Upstream Tracking and the Decay $B^0 \to K^+\pi^-\mu^+\mu^-$ at the LHCb Experiment

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Espen Eie Bowen

aus

United Kingdom

Promotionskomitee

Prof. Dr. Ulrich Straumann (Vorsitz)

Prof. Dr. Nicola Serra

Dr. Olaf Steinkamp

Dr. Barbara Storaci

Zürich, 2017

Abstract

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed to search for indirect evidence of New Physics in *CP* violation and rare decays of beauty and charm hadrons. The unique geometry takes advantage of the large b and c quark production in the forward region at the LHC. The detector includes a high granularity silicon-strip vertex detector, a silicon-strip detector upstream of the magnet and three stations of silicon-strip detectors and straw drift tubes downstream of the magnet.

This thesis is divided into two main parts. The first part details the development of improved algorithms to perform track reconstruction using the subdetectors upstream of the LHCb magnet. A novel idea to perform upstream tracking as an intermediate step of the track reconstruction sequence was investigated. The vast gains in tracking performance obtained when using upstream tracks led to the algorithm being adopted into the default reconstruction sequence for the LHCb Upgrade. It will play a large role in allowing LHCb to become the first hadron collider experiment to operate a software-only trigger at the full event rate. Following the success of upstream tracking for the Upgrade scenario, a similar strategy was developed for LHCb Run 2. The resulting algorithm was included in the tracking sequence in the first stage of the software trigger, greatly improving the signal efficiency for charm physics and allowing lifetime unbiased triggers for hadronic final states for the first time.

The second part of this thesis describes the measurements of the differential branching fraction and angular moments of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region. Proton-proton collision data are used, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$ collected by the LHCb experiment. Differential branching fraction measurements are reported in five bins of the invariant mass squared of the dimuon system, q^2, between 0.1 and $8.0\,\text{GeV}^2/c^4$. For the first time, an angular analysis that is sensitive to the S-, P- and D-wave contributions of this rare decay is performed. The set of 40 normalised angular moments describing the decay is presented for the q^2 range 1.1–$6.0\,\text{GeV}^2/c^4$.

Zusammenfassung

Der LHCb Detektor ist ein einarmiges Magnetspektrometer, welches den Pseudorapiditätsbereich $2 < \eta < 5$ abdeckt und für die Suche nach indirekten Anzeichen neuer Physik in *CP* verletzenden und seltenen Zerfällen optimiert ist. Die besondere Geometrie des Detektors nutzt den großen Wirkungsquerschnitt für die Produktion von b und c Quarks in der Vorwärtsrichtung am LHC besonders gut aus. Ein wichtiger Teil des Detektors dient der Rekonstruktion der Spuren und Impulse geladener Teilchen und besteht aus einem hochauflösenden Siliziumstreifen-Vertexdetektor, einem Siliziumstreifen-Spurdetektor vor dem Spektrometermagneten und drei Spurdetektoren aus Siliziumstreifen und Driftröhrchen hinter dem Magneten.

Diese Arbeit besteht aus zwei Hauptteilen. Der erste Teil der Arbeit behandelt die Entwicklung verbesserter Algorithmen zur Rekonstruktion der Trajektorien geladener Teilchen in den Spurdetektoren vor dem Magneten. Ein neuartiger Ansatz wurde entwickelt, der die Rekonstruktion von Spursegmenten vor dem Magneten als Zwischenschritt in der Spurrekonstruktion benutzt. Es wurde demonstriert, dass dieser Ansatz eine deutliche Verbesserung in der Leistungsstärke der Spurrekonstruktion ermöglicht, und dieser Ansatz wurde daraufhin als fester Bestandteil in die Rekonstruktionssequenz für den LHCb Upgrade integriert. Er wird einen wichtigen Beitrag dazu leisten, dass LHCb als erstes Experiment an einem Hadronen-Beschleunigerring in der Lage sein wird, die Datenselektion ausschliesslich auf Softwarealgorithmen beruhen zu lassen, welche die Daten mit der vollen Ereignisrate des Beschleunigers verarbeiten werden können. Aufbauend auf der erfolgreichen Entwicklung dieses Algorithmus für den LHCb Upgrade wurde eine analoge Strategie für die momentan laufende zweite Datennahmeperiode des LHCb Experiments entwickelt. Der resultierende Algorithmus wurde erfolreich in die Rekonstruktionssequenz im ersten Niveau der Software-Triggerselektion integriert, was zu einer signifikanten Verbesserung der Signaleffizienz für Charmphysik geführt hat und es zum ersten Mal ermöglicht hat, Selektionen für Zerfälle in rein hadronische Endzustände zu entwickeln, welche keine Verzerrung der Verteilung der Zerfallszeiten verursachen.

Der zweite Teil dieser Arbeit beschreibt die Messung des differenziellen Verzweigungsverhältnisses und der Momente der Winkelverteilungen für den Zerfall $B^0 \to K^+\pi^-\mu^+\mu^-$ in der Region der $K^*_{0,2}(1430)^0$-Resonanzen. Die Analyse

basiert auf den Daten von Proton-Proton Kollisionen im Umfang einer integrierten Luminosität von $3\,\mathrm{fb}^{-1}$, welche vom LHCb Experiment gesammelt wurden. Differenzielle Verzweigungsverhältnisse wurden in fünf Intervallen der quadrierten invarianten Masse q^2 des Zwei-Müonen-Systems zwischen 0.1 and $8.0\,\mathrm{GeV}^2/c^4$ bestimmt. Zum ersten Mal wurde eine Analyse der Winkelverteilungen in diesem selten Zerfall durchgeführt, welche auf Beiträge von S-, P- and D-Wellen sensitiv ist. Der komplette Satz der 40 normalisierten Momente, welche die Winkelverteilung in diesem Zerfall beschreiben, wurde für den q^2-Bereich 1.1–$6.0\,\mathrm{GeV}^2/c^4$ bestimmt.

Contents

1 Introduction

Particle physics is the study of the basic constituents of matter and their interactions, and aims to describe the fundamental laws governing the nature of the physical universe. It is a vast, diverse field encompassing both theoretical and experimental communities.

This thesis will focus on the LHCb experiment, situated at the Large Hadron Collider (LHC) at CERN. The LHC is the world's largest and most powerful particle accelerator and represents the forefront of experimental particle physics research. LHCb is one of several detector experiments designed to study the debris produced by the colliding beams of particles.

Chapter 2 provides a theoretical motivation for the experimental studies presented in later chapters. A brief introduction to the framework of modern particle physics is given, followed by a description of the properties of the decays of rare B mesons.

Chapter 3 describes the LHCb experiment, both current and future, with emphasis given to the tracking sub-detectors and trigger system.

Chapter 4 introduces track reconstruction in LHCb. A description of each of the tracking algorithms employed during LHCb Run 1 is provided, along with an explanation of how tracking performance is characterised.

Chapter 5 details the development of an improved upstream tracking algorithm for use in the reconstruction sequence of the LHCb Upgrade trigger. The motivation and initial performance are given, followed by a comprehensive description of the improvements made to the algorithm and the subsequent gains in performance.

Chapter 6 describes how the upstream tracking algorithm was subsequently adapted for use in the reconstruction sequence of the LHCb Run 2 trigger and the improvements in performance achieved.

Chapter 7 describes the measurements of the differential branching fraction and angular moments of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region. The dataset used has been collected by the LHCb experiment in pp collisions at centre of mass energies of 7 and 8 TeV, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$.

2 Theoretical overview

2.1 The Standard Model of particle physics

The Standard Model (SM) of particle physics incorporates the Glashow-Weinberg-Salam theory of the electroweak interaction and quantum chromodynamics. It describes our current understanding of elementary particles and their interactions. While this chapter briefly describes the most relevant aspects of the SM for this thesis, a more complete description can be found in Refs. [1, 2, 3].

Elementary particles can be divided into two categories: fermions, with half integer spin, and bosons, with integer spin. The SM fermions consist of quarks and leptons and are often referred to as matter. The properties of the twelve SM fermions are shown in Table 1. The SM bosons mediate the fundamental interactions between the fermions and amongst themselves. The properties of the SM bosons are shown in Table 2.

The SM is formulated within the framework of Quantum Field Theory (QFT) where the fermions and bosons are represented by quantum fields. The dynamics of the SM particles are described by the SM Lagrangian,

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} + \mathcal{L}_{Higgs}, \tag{2.1}$$

which is invariant under local transformations of the gauge group $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. The requirement of local gauge symmetry gives rise to gauge

Table 1: Properties of the quarks and leptons in the SM. The particle masses are taken from Ref. [4].

Quarks			**Leptons**		
Flavour	Mass [MeV/c^2]	Charge	Flavour	Mass [MeV/c^2]	Charge
u	$2.3^{+0.7}_{-0.5}$	$+2/3$	ν_e	$< 2 \times 10^{-6}$	0
d	$4.8^{+0.7}_{-0.3}$	$-1/3$	e	0.510998928(11)	-1
c	1275 ± 25	$+2/3$	ν_μ	< 0.19	0
s	95 ± 5	$-1/3$	μ	105.6583715(35)	-1
t	$(160^{+5}_{-4}) \times 10^3$	$+2/3$	ν_τ	< 18.2	0
b	4180 ± 30	$-1/3$	τ	1776.86(12)	-1

Table 2: Properties of the bosons in the SM. The particle masses are taken from Ref. [4].

Boson	Mass [GeV/c^2]	Charge	Spin
g	0	0	1
$W^{\pm}$	80.385 ± 0.015	± 1	1
Z	91.187 ± 0.0021	0	1
γ	0	0	1
H^0	125.09 ± 0.24	0	0

bosons corresponding to each of the symmetry groups. The $\mathrm{SU}(3)_C$ group introduces eight gluons, g, and represents the symmetry group of the strong interaction, resulting in the conservation of color charge, C. The $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ group introduces the $W^{\pm}$, Z and γ bosons, and represents the symmetry group of the electroweak interaction, resulting in the conservation of weak isospin, T, and weak hypercharge, Y.

The $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry is spontaneously broken by the Higgs field into the electromagnetic and weak interactions. This mechanism also generates the masses for the $W^{\pm}$ and Z bosons. The quarks and leptons obtain their masses via Yukawa couplings to the Higgs field. The fact that the Yukawa couplings are not diagonal generates the quark mixing described in Sec. 2.2. The Higgs boson, H^0, was the final particle of the SM to be discovered when it was confirmed in 2012 [5, 6].

2.2 Flavour violation in the SM

When only the three lightest quarks (u, d, s) were known, Cabibbo [7] proposed a mechanism to explain the observed rates of semileptonic processes. Vertices of the type $d \to u + W^-$ were given a factor $\cos\theta_C$, while those of the type $s \to u + W^-$ were given a factor $\sin\theta_C$, where $\theta_C \sim 13^\circ$ is the Cabibbo angle. This mechanism was found to be very successful in describing many decay rates but there was a problem: it allowed the K^0 meson to decay into a $\mu^+\mu^-$ pair. The amplitude of this process should be proportional to $\sin\theta_C \cos\theta_C$ but the calculated rate was much greater than the experimental limit.

A solution was proposed by Glashow, Iliopoulos and Maini (GIM) [8] who

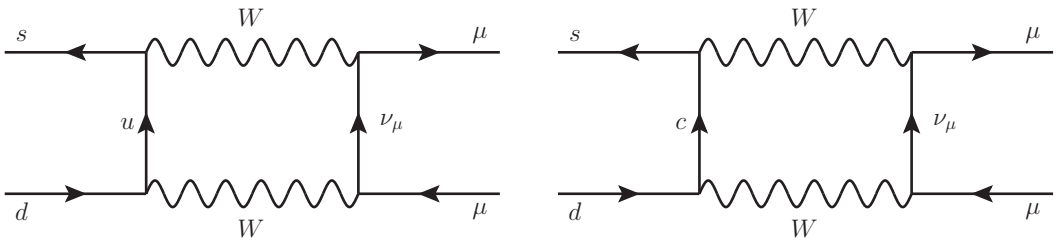


Figure 1: Feynman diagrams contributing to the decay $K^0 \to \mu^+\mu^-$. The amplitude for each diagram is proportional to the product of the coupling at each vertex: (a) $\sin\theta_C \cos\theta_C$, (b) $-\sin\theta_C \cos\theta_C$. The total amplitude is found by summing the two contributions.

introduced the fourth quark, c, with vertex couplings of $-\sin\theta_C$ and $\cos\theta_C$ for $d \to c + W^-$ and $s \to c + W^-$ transitions respectively. This was several years before the c quark was observed experimentally [9, 10]. According to the GIM mechanism, the diagram for the decay $K^0 \to \mu^+\mu^-$ in Fig. 1a is almost exactly cancelled by the corresponding diagram in Fig. 1b with a virtual c quark replacing the u. The small rate that remains is due to the differences in the masses of the u quark and the c quark.

In the Cabibbo-GIM scheme, it can be intepreted that the W does not couple to the mass eigenstates d and s but instead to the weak eigenstates d' and s', given by

$$d' = d\cos\theta_C + s\sin\theta_C, \quad s' = -d\sin\theta_C + s\cos\theta_C \tag{2.2}$$

or in 'Cabibbo matrix' form

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{2.3}$$

Even before the discovery of the c quark, Kobayaski and Maskawa [11] had generalised the Cabibbo-GIM scheme to incorporate three generations of quarks. This was motivated by the desire to explain CP violation within the Cabibbo-GIM scheme, which was not possible with only two generations. The weak eigenstates are related to the mass eigenstates through the CKM matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.4}$$

where each matrix element, V_{ij}, specifies the coupling of i to j. The additional generation of quarks increases the number of free parameters from one (θ_C) to four: three 'generalised Cabibbo angles' (θ_{12}, θ_{23}, θ_{13}) and a complex phase δ. This complex phase generates *CP* violation in the quark sector.

2.3 Beyond the SM

Despite its considerable success in predicting cross-sections and branching ratios of particles with great accuracy, the SM it is not without its limitations, some of which are outlined below

- ✦ Gravity, the fourth fundamental force, is not incorporated in the SM. However, this is not specific to the SM as there is no consensus on how to include gravity in a QFT.
- ✦ The SM has 19 free parameters: the 9 fermion masses, the 4 parameters of the CKM matrix, 3 gauge coupling constants, the Higgs vacuum expectation value v, the Higgs quartic coupling λ and the QCD θ parameter.
- ✦ Dark matter, which is believed to be five times as abundant as ordinary matter, is not accounted for in the SM.
- ✦ The amount of *CP* violation predicted by the SM is ten orders of magnitude lower than what is needed to account for the observed matter-antimatter asymmetry in the Universe, assuming matter and antimatter were created in equal amounts in the Big Bang.
- ✦ The measured mass of the Higgs boson implies a large cancellation between the bare mass and quantum corrections. The nature of a mechanism that could provide a natural cancellation of these corrections is unknown.

These limitations have led physicists to look for extensions to the SM in the form of New Physics (NP) models. Such NP models often predict new particles that may contribute to, and therefore modify, flavour-changing processes in the SM.

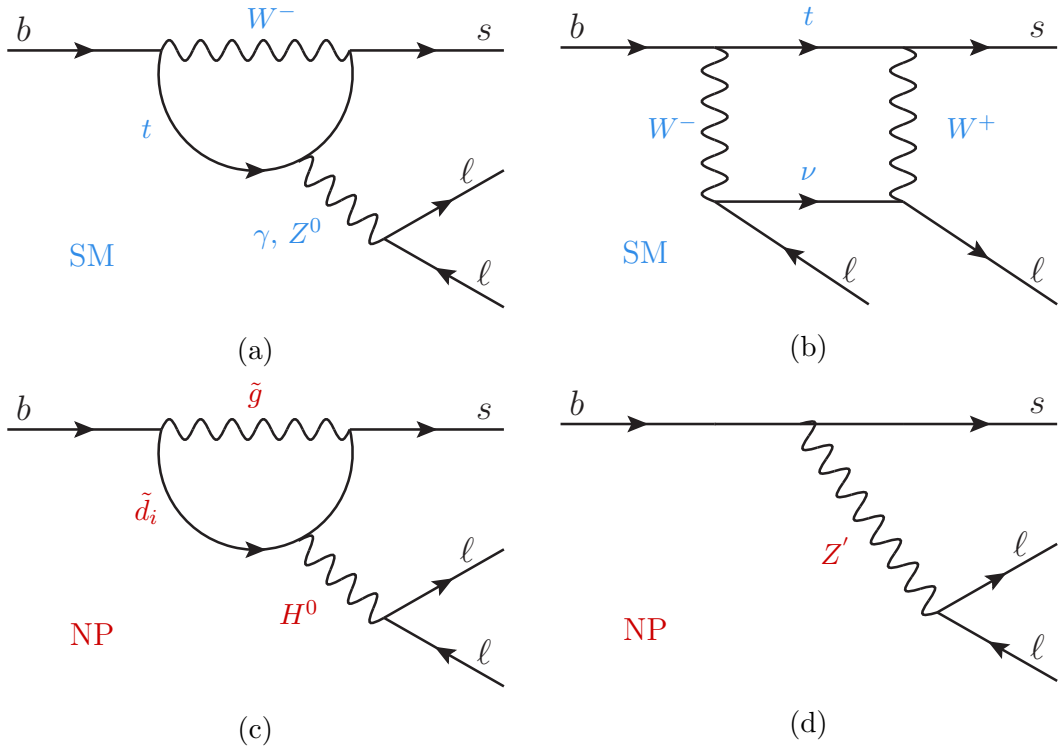
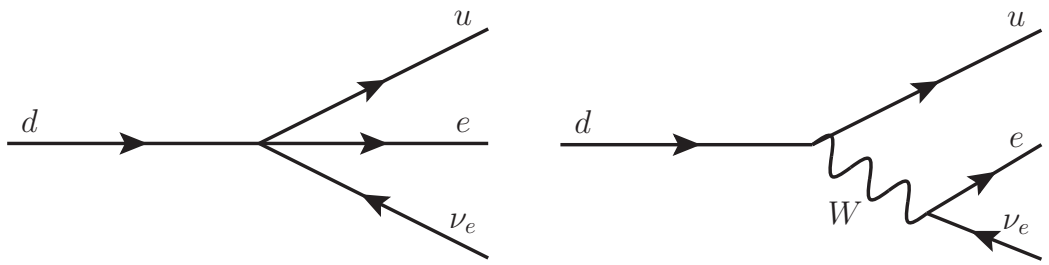


Figure 2: Feynman diagrams for the FCNC transition $b \to s\ell^+\ell^-$: in the SM (a,b) and in possible NP scenarios (c,d).

2.4 Rare decays of B mesons

Flavour-changing neutral current (FCNC) transitions are forbidden at tree level in the SM. They are rare processes that occur via loop and box diagrams, making them an ideal place to perform a search for NP. Rare decays of B mesons, such as those containing a $b \to s\ell^+\ell^-$ transition, can have sizeable NP contributions that are not swamped by the competing SM process.

By studying the properties of FCNC processes it is possible to perform model independent searches sensitive to a wide range of NP models. Feyman diagrams for the $b \to s\ell^+\ell^-$ transition, both in the SM and in possible NP scenarios [12, 13, 14], are shown Fig. 2. New particles can modify the dynamics of a decay: changing the branching fraction or the kinematic distribution of the final state particles. By comparing the SM prediction of observables with experimental measurements, the flavour structure of NP models can be probed. The theoretical framework used to interpret flavour physics measurements in a model independent way is the so-called Operator Product Expansion (OPE) [15].



(a) (b)

Figure 3: Feynman diagrams for β-decay in the (a) effective theory and (b) full theory.

In the OPE approach, all degrees of freedom above a given energy scale, Λ, are integrated out. This is valid as long as Λ is much larger than the energy scale of the decay, μ, which for the study of B mesons is chosen to be $\mathcal{O}(m_b)$. This formalism is analogous to Fermi's effective theory of weak decay in which the full theory is reduced to a four point interaction, as shown in Fig. 3 for β-decay.

In the SM, the effective Hamiltonian describing $b \to s\ell^+\ell^-$ decays can be written as

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \sum_i [\mathcal{C}_i(\mu)\mathcal{O}_i(\mu) + \mathcal{C}_i'(\mu)\mathcal{O}_i'(\mu)]. \tag{2.5}$$

where G_F is the Fermi constant and V_{ij} are CKM matrix elements. The complex Wilson coefficients, $\mathcal{C}_i$, incorporate the short distance (high energy) contributions. For each Wilson coefficient, there is a local operator $\mathcal{O}_i$ which incorporates the long distance (low energy) contributions. The primed operators represent the right-handed currents, which are highly suppressed in the SM. The advantage of the OPE approach is that the Wilson coefficients are independent of the underlying process and can include contributions from NP.

For $b \to s\ell^+\ell^-$ decays, the dominant contributions in the SM arise from the following operators [16]

$$
\begin{aligned}
\mathcal{O}_7 &= \frac{e}{g^2} m_b (\overline{s}\sigma_{\mu\nu} P_R b) F^{\mu\nu}, & \mathcal{O}'_7 &= \frac{e}{g^2} m_b (\overline{s}\sigma_{\mu\nu} P_L b) F^{\mu\nu}, \\
\mathcal{O}_9 &= \frac{e^2}{g^2} (\overline{s}\gamma_\mu P_L b)(\overline{\ell}\gamma^\mu \ell), & \mathcal{O}'_9 &= \frac{e^2}{g^2} (\overline{s}\gamma_\mu P_R b)(\overline{\ell}\gamma^\mu \ell), \\
\mathcal{O}_{10} &= \frac{e^2}{g^2} (\overline{s}\gamma_\mu P_L b)(\overline{\ell}\gamma^\mu \gamma_5 \ell), & \mathcal{O}'_{10} &= \frac{e^2}{g^2} (\overline{s}\gamma_\mu P_R b)(\overline{\ell}\gamma^\mu \gamma_5 \ell), \\
\mathcal{O}_S &= \frac{e^2}{16\pi^2} m_b (\overline{s} P_R b)(\overline{\ell}\ell), & \mathcal{O}'_S &= \frac{e^2}{16\pi^2} m_b (\overline{s} P_L b)(\overline{\ell}\ell), \\
\mathcal{O}_P &= \frac{e^2}{16\pi^2} m_b (\overline{s} P_R b)(\overline{\ell}\gamma_5 \ell), & \mathcal{O}'_P &= \frac{e^2}{16\pi^2} m_b (\overline{s} P_L b)(\overline{\ell}\gamma_5 \ell),
\end{aligned}
\tag{2.6}
$$

where $P_{L/R} = (1 \mp \gamma_5)/2$ are the left- and right-handed projection operators. The operator $\mathcal{O}_7$ is the electromagnetic operator corresponding to the emission of a photon. The vector and axial-vector operators, $\mathcal{O}_9$ and $\mathcal{O}_{10}$, describe the Z penguin and W box diagrams. The operators $\mathcal{O}_S$ and $\mathcal{O}_P$ represent scalar and pseudoscalar operators.

3 The LHCb experiment

3.1 LHC

The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) is a particle accelerator designed to collide protons at a center-of-mass energy ($\sqrt{s}$) of 14 TeV [17]. It has a peak design (instantaneous) luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for proton operation. It is also used to accelerate heavy ions. The accelerator complex at CERN is shown in Fig. 4. The four main experiments are ATLAS, CMS, LHCb and ALICE. ATLAS and CMS are both multipurpose experiments, while ALICE is used for the study of heavy ion collisions. The present thesis concerns the LHCb experiment, of which an overview of the goals and design of the experiment is given below.

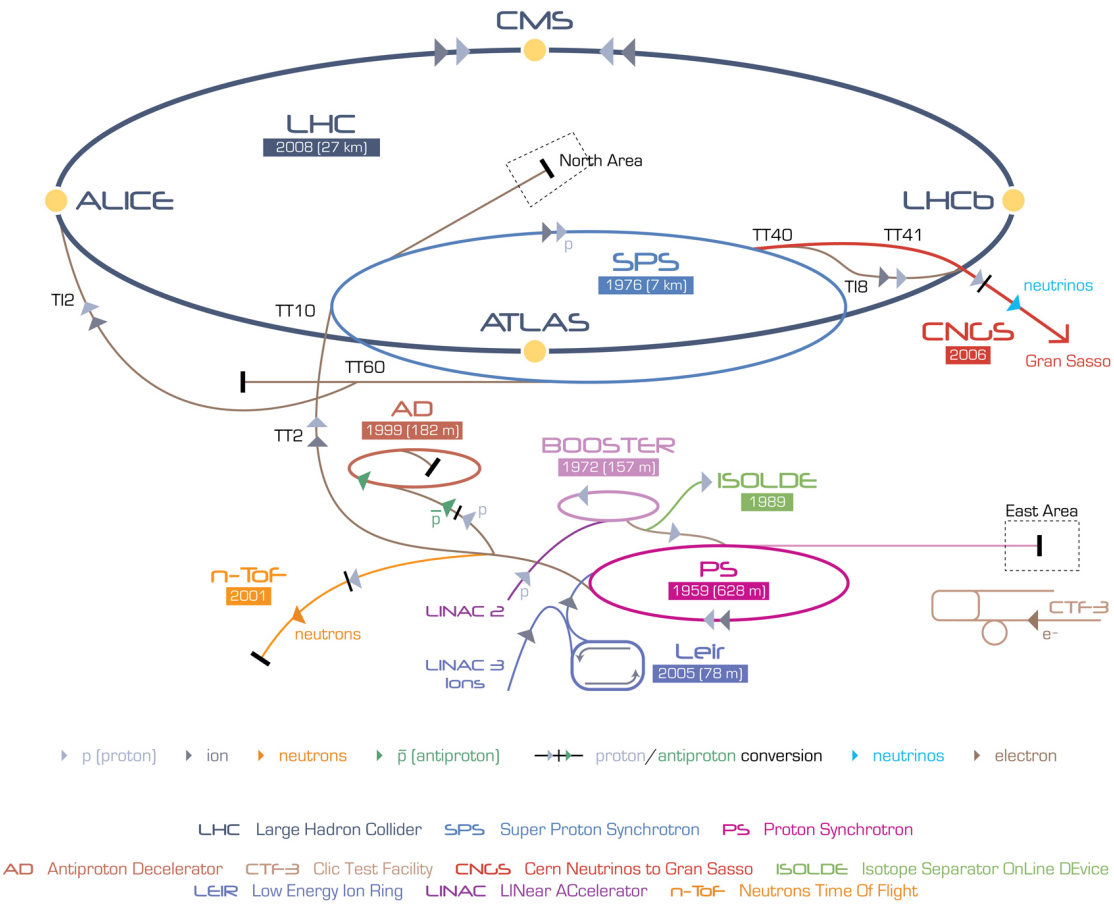


Figure 4: Schematic view of the CERN accelerator complex.

3.2 LHCb experiment

LHCb is a dedicated heavy flavour experiment with the primary goal of searching for indirect evidence of NP in *CP* violation and rare decays of beauty and charm hadrons [18]. During Run 1 of the LHC (2010-2013), the LHCb experiment took data at center-of-mass energies of $\sqrt{s}$ = 7-8 TeV. Due to the large beauty and charm production cross-sections provided by the LHC, the LHCb experiment was able to collect $\sim 10^{12}$ heavy flavour decays during data taking in 2011 and 2012 [19].

The data used for physics analysis in this thesis was collected in 2011 and 2012. In 2011, the center-of-mass energy was 7 TeV and the majority of the data was taken at a instantaneous luminosity of $3.5 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. In 2012, the center-of-mass energy was increased to 8 TeV and the data was taken at a instantaneous luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This corresponds to an integrated luminosity of $1.11\,\mathrm{fb}^{-1}$ in 2011 and $2.08\,\mathrm{fb}^{-1}$ in 2012.

The LHCb detector is a single arm forward spectrometer that covers a pseudo-rapidity range of $2 < \eta < 5$. The layout of the LHCb detector, shown schematically in Fig. 5, is motivated by the physics of b quark production at the LHC. The dominant production processes for $b\bar{b}$ pairs in pp collisions are gluon-gluon fusion and quark-antiquark annihilation, as illustrated in Fig. 6. In the high energy collisions present at the LHC, the $b\bar{b}$ pairs tend to be produced in the same forward or backward cone, shown in Fig. 7. Therefore, due to its unique geometry, LHCb is able to detect $\sim$ 40% of the heavy quark pairs despite covering only $\sim$ 4% of the total solid angle.

The following requirements are crucial in order to fulfil the LHCb physics programme: efficient, robust and flexible triggering on a variety of different final states, excellent tracking (momentum, impact parameter (IP) and primary vertex (PV) resolution), precise decay time resolution and excellent particle identification. The importance of each of these requirements are discussed in detail in the following sections.

3.2.1 Dipole magnet

A warm dipole magnet [21] is used in order to allow a measurement of the momentum of charged particles in the forward acceptance of ± 250 mrad vertically and ± 300 mrad horizontally. It provides an integrated magnetic field of 4 Tm for tracks of 10 m in length. The strength of the main component of the magnetic

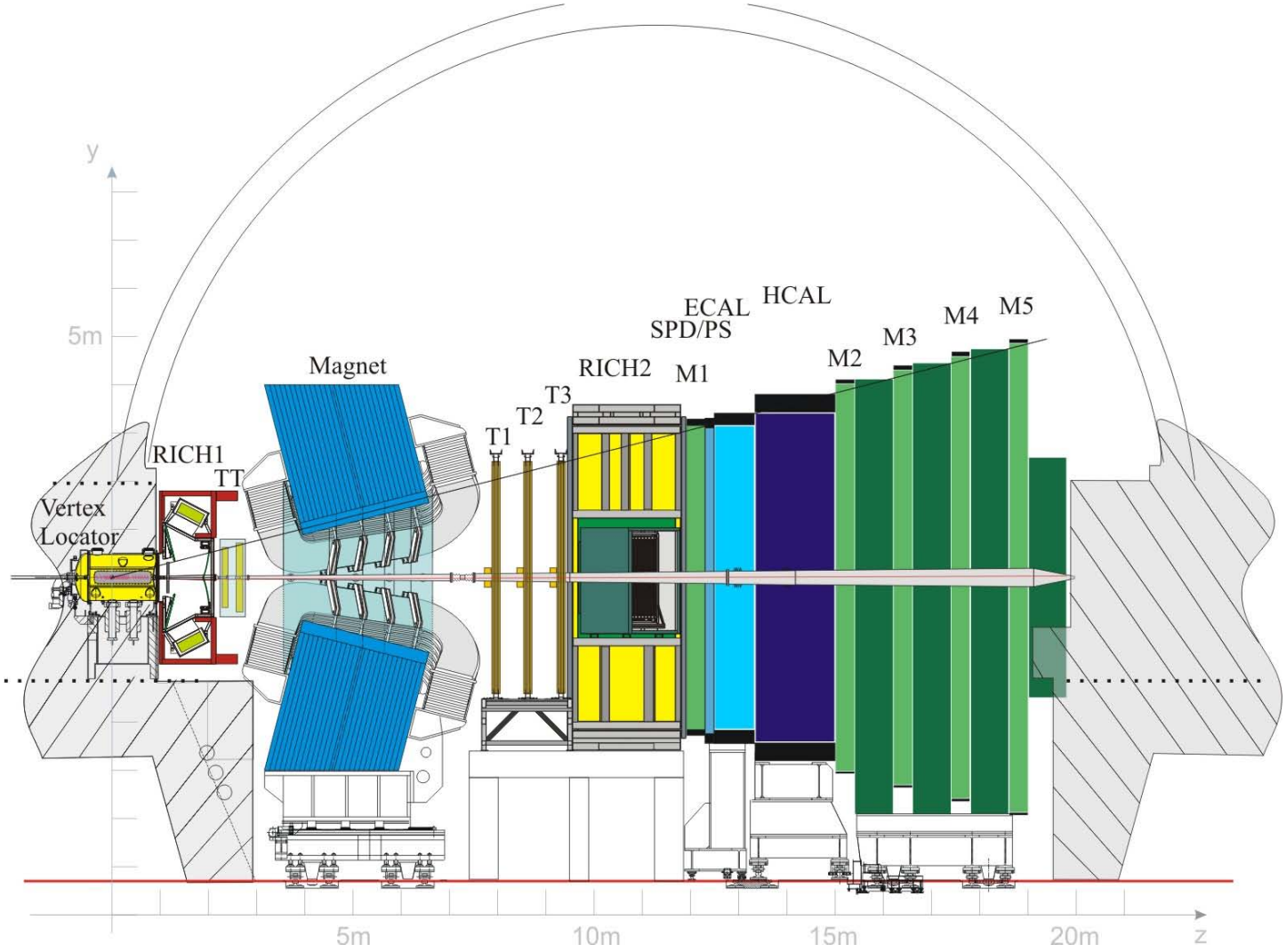
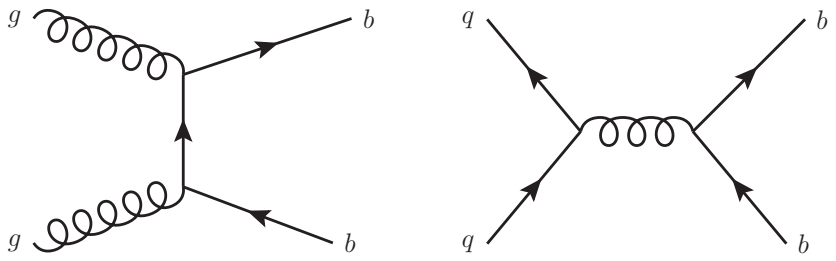


Figure 5: Schematic view of the LHCb detector.



(a) (b)

Figure 6: Feynman diagrams of the dominant production processes for $b\bar{b}$ pairs in proton-proton collisions. These are (a) gluon-gluon fusion and (b) quark-antiquark annihilation [20].

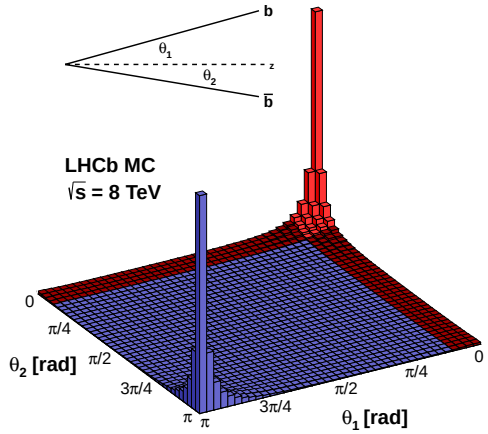


Figure 7: The angular distribution of $b\bar{b}$ pairs in terms of the polar angle from the beam axis. The red area shows the acceptance of the LHCb detector.

field, B_y, as a function of z is shown in Fig. 8. The location of the tracking sub-detectors, described in Sec. 3.2.2, are also shown. The polarity of the magnetic field is changed regularly in order to be able to control possible detection asymmetries.

3.2.2 Tracking system

The trajectory of charged particles traversing the LHCb detector are reconstructed using a dedicated tracking system. The tracking system consists of a high granularity vertex detector surrounding the pp interaction region, a large area silicon-strip

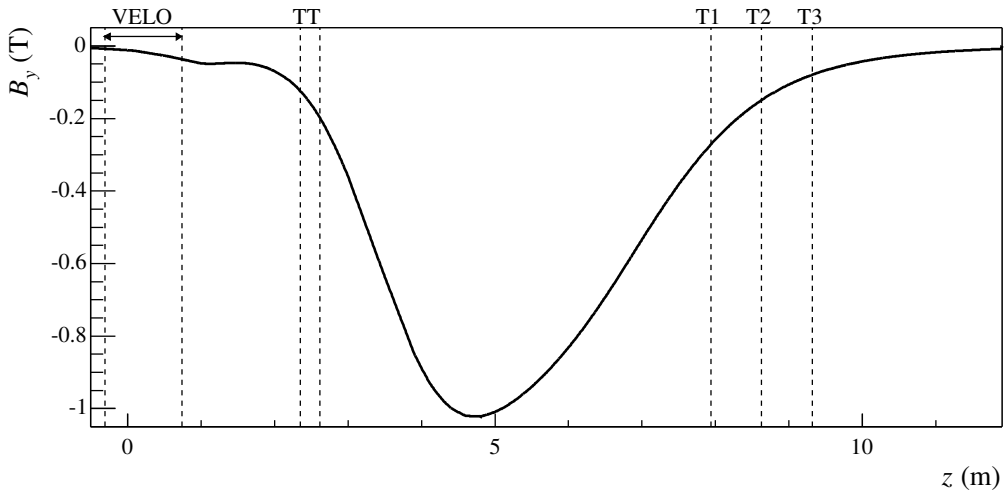


Figure 8: The strength of the main component of the magnetic field, B_y, as a function of z. The location of the tracking sub-detectors are shown.

detector located upstream of the magnet, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. By determining the deflection of the charged particles that have passed through the magnetic field their momentum can be measured.

The Vertex Locator (VELO) [22, 23] is a silicon microstrip detector that provides measurements of track coordinates close to the pp interaction region. These precision measurements are used to identify both the primary interaction vertices and the displaced secondary vertices of the decays of b and c hadrons. The detector consists of 42 silicon micro-strip stations with r-ϕ geometry, shown schematically in Fig. 9. It has two retractable halves which are opened and closed during each LHC fill. When closed, the sensors are positioned only 7 mm from the LHC beam. This is essential in achieving precise IP measurements.

The Silicon Tracker (ST) consists of two silicon microstrip detectors: the Tracker Turicensis[1] (TT) [24] and the Inner Tracker (IT) [25]. The TT is located upstream of the magnet and covers the full acceptance of the experiment. The IT is located downstream of the magnet and covers a cross-shaped region at the center of the three tracking stations. Each ST station contains four layers with a x-u-v-x layout such that the u and v layers are tilted by $\pm$ 5° with respect to the vertical. The inclined layers allow stereo measurements to be made. A schematic

[1]Formerly known as the Trigger Tracker.

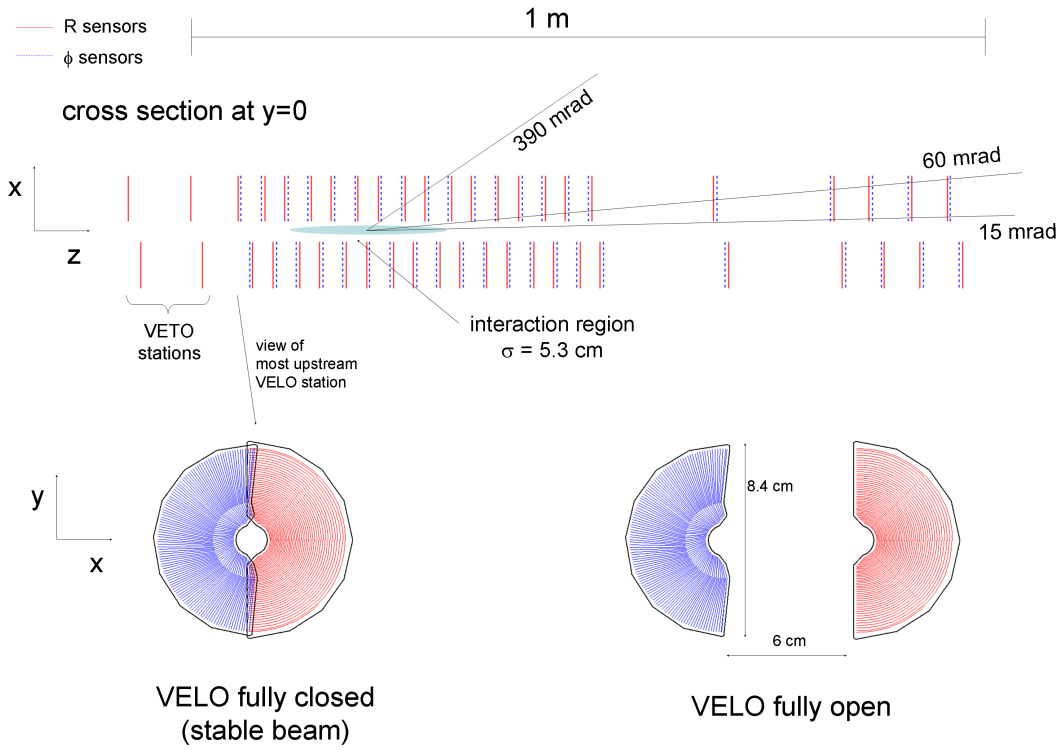


Figure 9: Schematic view of the VELO detector.

view of the TT sub-detector is shown in Fig. 10. As the beampipe passes through the TT there is a square hole at the center of the detector which reduces its acceptance. The square hole has a width of 7.7 cm at the first TT sub-station (TTa) and 8.0 cm at the second (TTb).

The Outer Tracker (OT) [26, 27, 19] is a drift-time detector made up of arrays of individual straw-tube modules. The modules each contain two staggered layers of drift-tubes of 4.9 mm in diameter. They are filled with a mixture of Argon (70%), CO_2 (28.5%) and O_2 (1.5%), which provides both fast drift times and sufficient drift-coordinate resolution. The detector consists of three stations each with four layers arranged with a x-u-v-x layout such that the u and v layers are tilted by $\pm$ 5° with respect to the vertical.

The track reconstruction efficiency can be defined as the probability that the trajectory of a charged particle that has passed through the full tracking system is reconstructed. This can be measured in data using a tag-and-probe method with $J/\psi \to \mu^+\mu^-$ decays [28]. One of the daughters is fully reconstructed (tag), while

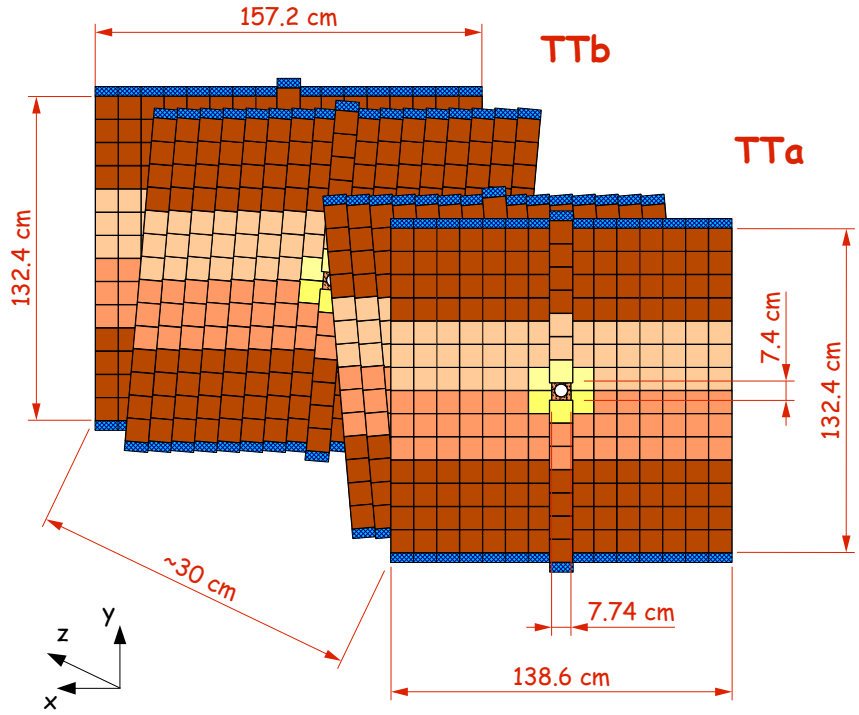


Figure 10: Schematic view of the TT sub-detector geometry. The different colours indicate the number of sensors that are wire bonded together in the y direction.

the other is only partially (probe), although well enough to reconstruct the J/ψ invariant mass. The efficiency can then be measured by matching the probe track to a fully reconstructed track. The average efficiency is found to be over 95% and is only slightly affected in high multiplicity events [28].

The momentum resolution for tracks passing through the full tracking system can be measured in data using $J/\psi \to \mu^+\mu^-$ decays. The relative momentum resolution, $\delta p/p$ is found to be between 0.4% - 0.6% for tracks up to 100 GeV/c [19]. The mass resolution is determined from data by studying the J/ψ, $\psi(2S)$, Υ and Z resonances. The relative mass resolution, σ_m/m, is found to be about 5 per mille up to the Υ masses [19].

Precise vertex resolution is important to allow the separation of primary and secondary decay vertices. The primary vertex resolution depends strongly on the number of tracks used to form it. It can be measured in data in an event-by-event manner by randomly splitting the track sample in two and reconstructing the PV using each independent set of tracks. In 2011 data, a 25-track vertex was found

to have a resolution of 13 μm in x and y and 71 μm in z [19].

While the reconstructed decay time of charm and beauty hadrons is used in offline selections and for precise measurements of lifetimes, the most stringent requirement on the decay time resolution originates from the need to resolve the fast B_s^0-$\overline{B}_s^0$ oscillations in mixing. The decay time resolution is topology dependent and is calibrated in data for each final state using prompt combinations that fake the signal candidates. The shape of the prompt decay time distribution is determined only by the resolution function. The typical decay time resolution is 45 fs for a 4-track vertex [19].

3.2.3 Particle identification

Excellent particle identification (PID) is a crucial requirement for LHCb. Charged particle identfication is important to be able distinguish specific final states from those with otherwise identical topologies and to perform b quark flavour tagging. Detecting photons is essential to allow the reconstruction of rare radiative decays. As muons are present in the final state of many CP-sensitive decays, such as $B_s^0 \to J/\psi\,\phi$, and rare decays, such as $B_s^0 \to \mu^+\mu^-$, their triggering and identification are of fundamental importance.

The LHCb detector has several specific systems in order to perform the various particle identification tasks. The primary goal of the RICH system [24, 29] is the identification of charged hadrons (π, K, p). This is achieved by exploiting the Cherenkov effect by which a charged particle travelling in a medium will emit Cherenkov radiation whenever the velocity of the particle exceeds the velocity of light in that medium. The cone of light will be emitted at a particular angle (θ_c) relative to the particle path. RICH detectors use a combination of spherical and flat mirrors to focus the cone of light into a ring. This ring is then projected onto an array of photodetectors. The Cherenkov angle (θ_c) can be determined from the radius of the ring. The velocity of the particle is found from θ_c.[2] Combining this velocity with the momentum measured by the tracking systems allows the determination of the particle mass and, therefore, the type of particle.

In the forward region covered by the LHCb experiment, there is a strong anticorrelation between polar angle and momentum. Due to this, the RICH system consists of two detectors. The detector upstream of the magnet, RICH1, covers

[2] $\cos\theta_c = \frac{1}{n\beta}$ where $\beta = \frac{v}{c}$ and n is the refractive index of the medium

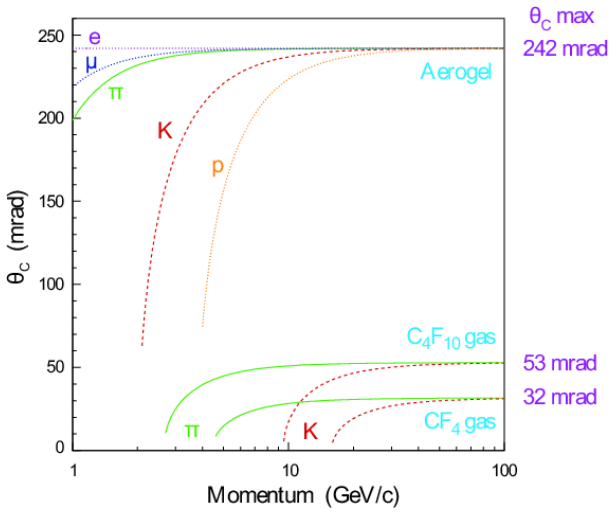


Figure 11: Cherenkov angle versus particle momentum for the three gas radiators used in the RICH system.

the low and intermediate momentum region 2-40 GeV/c for the full angular acceptance 15-300 mrad. The downstream detector, RICH2, covers the high momentum region 15-100 GeV/c over a reduced angular acceptance 15-120 mrad. During Run 1, RICH1 contained aerogel and C_4F_{10} as gas radiators while RICH2 contained CF_4.[3] The relationship between Cherekov angle and particle momentum is shown in Fig. 11 for each of the three radiators.

The calorimeter system [30] is used both to identify and to measure the position and energy of photons, electrons and hadrons. It consists of a Scintillating Pad Detector (SPD), a Preshower (PS), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). The ECAL is scintillator/lead sampling calorimeter with an energy resolution of $\sigma_E/\sqrt{E} = 1\% + 10\%/\sqrt{E}$. The HCAL is scintillator/iron sampling calorimeter with an energy resolution of approximately $\sigma_E/\sqrt{E} \sim 70\%/\sqrt{E}$.

The muon system [31, 32, 33, 34] is used to identify muons and consists of five stations. The first station, M1, is located upstream of the calorimeters and uses triple Gas Electron Multiplier (GEM) detectors. The remaining four stations, M2-M5, located downstream of the calorimeters use Multiwire Proportional Chambers (MWPC) which are interleaved with 80 cm thick iron absorbers to select penetrating muons.

For a given particle hypothesis (K, π, μ, p), a combined, overall likelihood

[3]The aerogel was removed during Long Shutdown 1 (2013-2015).

is obtained using information from the RICH detectors, the calorimeters and the muon system. The variable commonly used for PID selection requirements is the delta log likelihood (DLL), for example,

$$\mathrm{DLL}_{K\pi} = \log(\mathcal{L}_K) - \log(\mathcal{L}_\pi) \tag{3.1}$$

where $\mathcal{L}_K$ is the likelihood that the particle is a kaon and $\mathcal{L}_\pi$ is the likelihood that the particle is a pion.

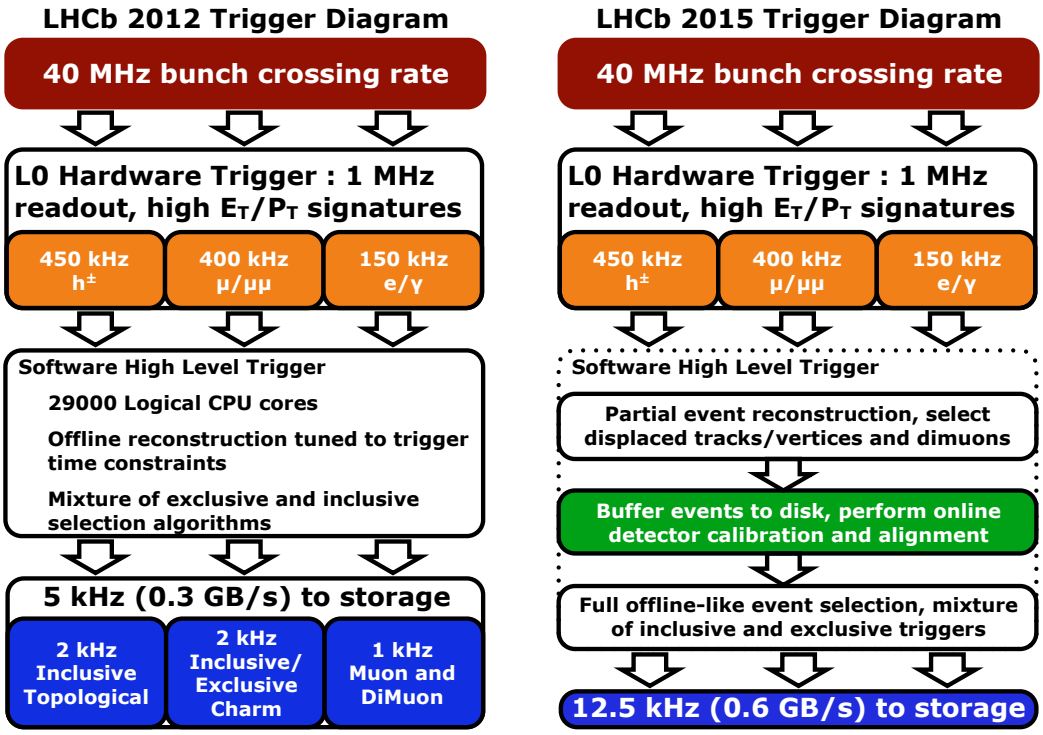
3.2.4 Trigger system

The LHCb trigger system [35, 36] plays an important role in selecting signal events and rejecting background. Two key signatures of the decays of beauty and charm hadrons are large invariant masses and large lifetimes with respect to light unflavoured particles. The large invariant masses result in the daughter particles having significant transverse momentum with respect to the beam axis (p_T). The large lifetimes lead to the daughter particles having a large IP with respect to the primary vertex. Several key channels also contain muons in the final state.

The trigger system consists of two stages: a hardware trigger (L0) followed by a high-level trigger implemented in software (HLT). The trigger scheme used in 2012 data taking is shown in Fig. 12a. The L0 trigger reduces the event rate from the rate of visible interactions at ∼13 MHz to 1 MHz at which the LHCb detector can be read out. Therefore, a decision based on information from the muon systems and calorimeters needs to be reached in less than 4 µs. Events that contain high p_T muons or large transverse energy deposits in either calorimeter are selected by L0.

In the first stage of the software trigger (HLT1), a partial event reconstruction is performed due to the limitations of the available computing power. The HLT is implemented in a online computing farm containing around 30 000 logical cores, giving it approximately 30 ms to process each event [37]. In order to reduce the execution time of the reconstruction sequence, only VELO track segments that either have a large IP with respect to the primary vertex or are matched to hits in the muon stations are extrapolated into the main tracking system.[4] The event is selected if a good quality track with a large p_T is found, further reducing the

[4]It should be noted that the IP requirements bias the observable lifetime distributions for fully hadronic b hadron decays.



(a) (b)

Figure 12: The trigger schemes used for (a) 2012 and (b) 2015 data taking.

event rate to 70 kHz.

Events selected by HLT1 are passed to the second stage of the software trigger (HLT2). In this stage, all tracks with a minimum p_{T} greater than 300 MeV/c are reconstructed without any requirement on IP or matched muon hits. A combination of exclusive and inclusive selections are used to reduce the event rate to 5 kHz, which is written to disk. The inclusive selection of heavy flavour decays with hadrons in the final state is performed by a "topological" algorithm containing a multivariate classifier that identifies b hadron decays with two-, three- and four-track vertices [38, 39].

3.3 LHCb Run 2

During Run 2 of the LHC (2015-2018), the LHCb experiment will take data at an increased center-of-mass energy of $\sqrt{s} = 13$ TeV. The bunch spacing will change from 50 ns to the design value of 25 ns. With these conditions the same instantaneous luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ as for 2012 data taking can be achieved with

a lower pile-up. The online computing farm will also be upgraded to include a 5 PB disk buffer.

The software trigger used in Run 1 contained only a simplified track reconstruction sequence with respect to the one used in the full offline event reconstruction. Furthermore, only a preliminary alignment and calibration of the detector was available and no information from the RICH system was used. For Run 2, the trigger system was redesigned with two key objectives: to enable the full offline event reconstruction to be performed within the trigger and to achieve the same alignment and calibration quality at the trigger level as could be achieved offline during Run 1 [40]. These two features allow physics analyses to be performed directly on the output of the software trigger [41].

The trigger scheme used in 2015 data taking is shown in Fig. 12b. Following the hardware stage and a partial event reconstruction, selected events are buffered to disk. The automatic alignment and calibration procedure is then performed [42]. Once the detector is aligned and calibrated, the full offline-like event reconstruction is performed. With the full information available, including PID from the RICH system, a wide range of inclusive and exclusive selections can be performed in order to trigger the event.

With the offline quality event reconstruction achieved in the new trigger scheme, it is also possible to only write out the information of the signal candidates [41]. This leads to a large saving in storage space ($\sim 90\%$) and is ideal for the analysis of channels with high yields that would previously have been heavily pre-scaled. It also allows rapid turn-around from data taking to analysis on the order of a few weeks [43, 44]. However, this trigger scheme requires careful planning of the selection criteria in order to not leave out potential interesting physics cases.

3.4 LHCb Upgrade

The LHCb experiment will undergo an upgrade during Long Shutdown 2 (2018-2019) to allow data taking at $\sqrt{s} = 14\,\text{TeV}$ with an instantaneous luminosity of $2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ [45, 46]. The upgraded detector will collect at least $5\,\text{fb}^{-1}$ a year over 10 years of operation.

Two key features of the LHCb Upgrade are the following: a trigger-less readout system and a full software trigger [47]. With the current experiment the collision

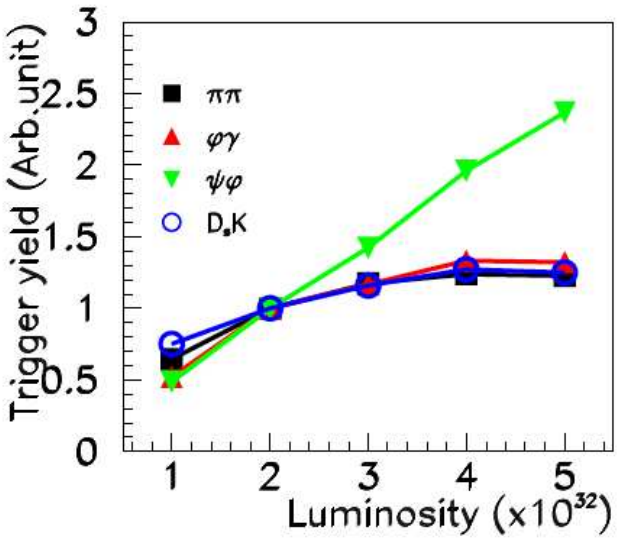


Figure 13: The trigger yield as a function of instantaneous luminosity for different decays of B mesons [45].

rate must be reduced to the readout rate of 1 MHz. This is achieved using basic information from the calorimeters and muon system as described in Sec. 3.2.4. The largest inefficiencies in the trigger chain, especially for purely hadronic decays, occur at this stage. Furthermore, this constraint inhibits the operation of the detector at higher instantaneous luminosities as trigger yields for hadronic channels would saturate, as shown in Fig. 13. Removing this bottleneck by implementing a trigger-less readout system will allow the full rate of visible interactions to be processed by a purely software trigger. This trigger will offer great flexibility and improve the trigger efficiency significantly for a number of physics channels. It will also allow lifetime unbiased hadronic triggering for the first time at a hadron collider.

In order to incorporate a trigger-less readout, all the front-end electronics need to be replaced. Futhermore, the upgraded detector will need to be able to cope with the factor five increase in instantaneous luminosity. Therefore, many of the existing sub-detectors will be replaced. The current VELO will be replaced by the VELO Pixel (VP) detector [48]. The VP will contain 41 million $55 \times 55\,\mu\text{m}$ pixel sensors with micro-channel CO_2 cooling. When closed, the innermost pixel will be located just 5.1 mm from the LHC beam. The TT will be replaced by the Upstream Tracker (UT) [49], shown schematically in Fig. 14. The UT will be a

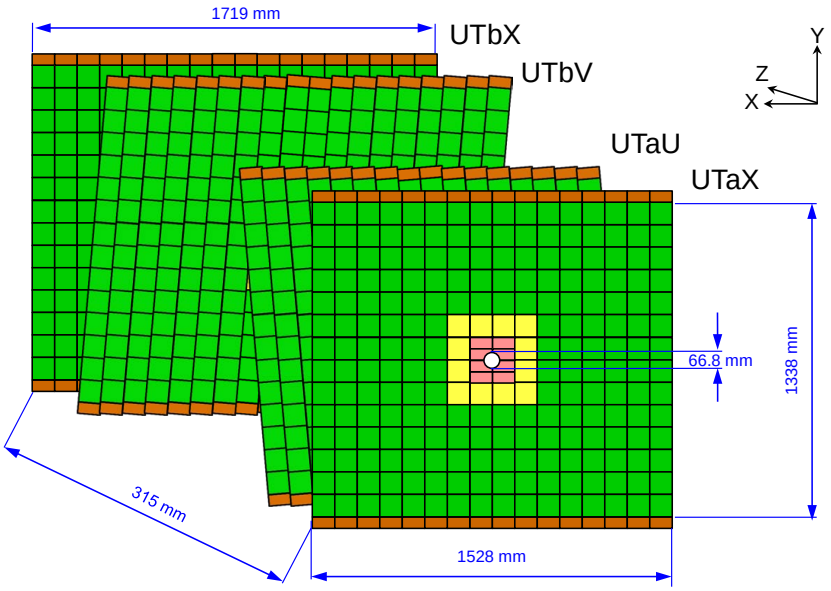


Figure 14: Schematic view of the UT sub-detector geometry. The different sensor geometries are colour coded.

high granularity silicon microstrip detector with improved coverage of the LHCb acceptance. The IT and OT will be replaced by the Scintillating Fibre Tracker (SciFi) [49]. This single, fast detector will contain 2.5 m long multilayer ribbons of 250 µm diameter scintillating fibres with silicon photomultiplier readout.

4 Track reconstruction in Run 1

4.1 Introduction

The track reconstruction in LHCb is performed by several different algorithms [50]. In order to describe the process, it is first necessary to introduce the notion of track types and track states which are described in Sec. 4.2 and Sec. 4.3 respectively. Each of the tracking algorithms used in Run 1 are described in detail in Sec. 4.4 with special emphasis given to the upstream tracking algorithm. The duplicate track removal and track fit procedures are described in Sec. 4.5 and Sec. 4.6 respectively. Finally, the methods used to determine the performance of the track reconstruction using simulation are detailed in Sec. 4.7.

4.2 Track types

The tracks reconstructed in the LHCb detector are divided into types depending on the sub-detectors in which they are reconstructed, as shown in Fig. 15. VELO tracks are defined as those which have measurements only in the VELO sub-detector. These tracks can be either forward or backward. Upstream tracks are defined as those which have measurements only in the VELO and TT (UT) sub-detectors. Upstream tracks are also referred to as VeloTT (VeloUT) tracks. T tracks are defined as those which have measurements solely in the T stations. Downstream tracks have measurements in the TT (UT) sub-detector and T stations. Long tracks have measurements in the VELO sub-detector and T stations and may also have measurements in the TT (UT) sub-detector. These tracks provide the best momentum resolution for particles which traverse the full tracking detector and are used in the majority of LHCb analyses.

4.3 Track states

In LHCb, a track is modelled as a series of straight line segments called track states. A track state is defined by a state vector of the form

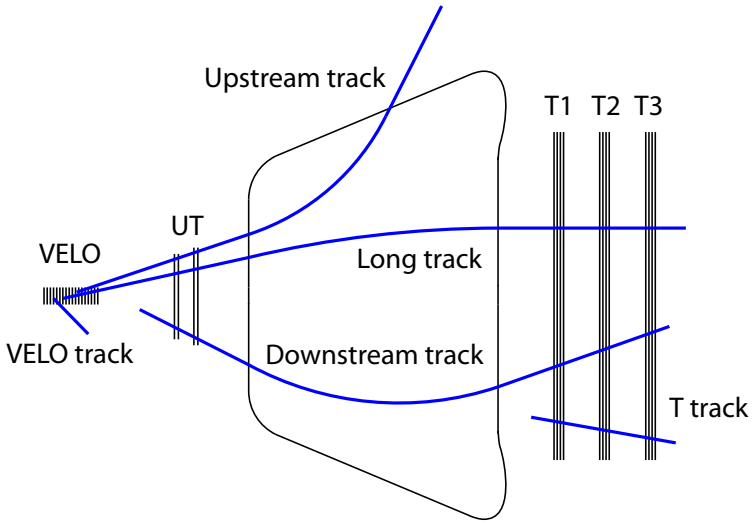


Figure 15: Schematic diagram of the LHCb tracking system. The various track types reconstructed by the different tracking algorithms are shown.

$$\vec{x} = \begin{pmatrix} x \\ y \\ t_x \\ t_y \\ q/p \end{pmatrix} \text{ with } t_x = \frac{\partial x}{\partial z} \text{ and } t_y = \frac{\partial y}{\partial z} \tag{4.1}$$

and a corresponding 5×5 state covariance matrix at a given position in z. Here, q and p are the charge and momentum of the track respectively.

4.4 Track reconstruction algorithms

In order to reconstruct the different track types, several tracking algorithms are employed. The two stand-alone algorithms, VELO tracking and track seeding, are described in Sec. 4.4.1 and Sec. 4.4.3 respectively. The other algorithms use input from these two algorithms in order to perform a further track reconstruction.

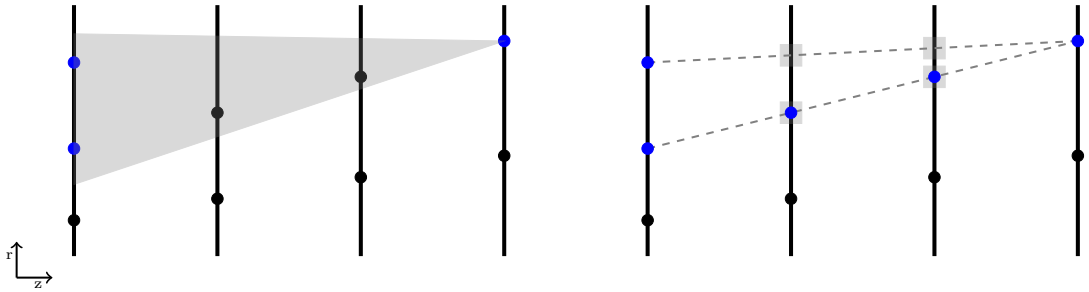


Figure 16: Quadruplets of hits are searched for in the VELO starting from the most downstream layer. Starting with a hit in that layer, a search window is opened in the fourth most downstream layer. From hits found within the window, the expected position of hits in the intermediate layers can be predicted assuming the track is a straight line in the r-z projection. If hits fall within a tolerance of the expected positions, quadruplets are formed and a track is created.

4.4.1 VELO tracking

The VELO tracking algorithm [51] is used to find tracks in the VELO. As there is a neglible magnetic field in the VELO, tracks are expected to be approximately straight lines. The track search begins in the most downstream layer of the VELO. Quadruplets of hits are searched for in the r-sensors as shown in Fig. 16. If they are found, they are extended back to smaller z adding hits that are consistent with coming from the same track. Next, the same quadruplet search is performed for backward-going tracks. Triplets are then searched for, first backward-going and then forward-going, requiring that the hits have not been used in the quadruplet search.

Starting from the longest r-z track, ϕ hits are searched for that are consistent with coming from the same track. These 3D tracks are then fitted with a χ^2 minimisation.

4.4.2 Forward tracking

The Forward tracking algorithm [52] is used to find long tracks. A Hough transform is utilised to associate hits in the T-stations to each VELO track. The VELO track is linearly extrapolated to the T-stations and a symmetric search window is opened in each x layer. The VELO track state and knowledge of the $\vec{B}$ field are used to project each selected hit to the z position of a reference plane. Hits from the same particle are expected to be projected to the same x position while

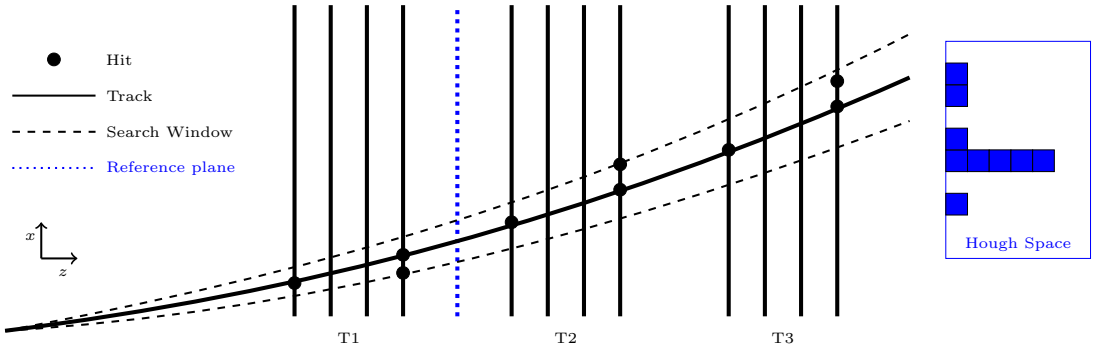


Figure 17: A Hough transform is used to associate hits in the T-stations to a VELO track. Each hit within a search window around the extrapolated track is projected to the z position of a reference plane. Hits from the same particle are expected to be projected to the same x position while random hits should be uniformly distributed.

random hits should be uniformly distributed. This procedure is shown schematically in Fig. 17. The resulting clusters are fitted and outliers are removed using a χ^2 criterium. An additional cluster search is used to add stereo hits that are consistent with the x-z track. This 3D track is then fitted, outliers are removed and the best track candidate is chosen based on its χ^2/dof.

4.4.3 T seeding

The T seeding algorithm [53] is used to find T tracks. Track candidates are first searched for in the x-z projection. A straight line is formed between suitable pairs of hits in T1 and T3. A compatible hit in T2 is added to form a parabola. Further hits in x layers consistent with this parabola are added to the track candidate. A Hough transform is used to add stereo hits. A weighted least-squares fit is then applied to each candidate.

4.4.4 Track matching

The track matching algorithm [54] is also used to form long tracks. It takes both VELO and T tracks as input (seeds). The difference in x and y of the two seeds are calculated by extrapolating them both to the magnet bending plane (Δx) and the end of the T-stations (Δy) respectively. A matching criterion χ^2 is formed using Δx, Δy, Δt_x and Δt_y. If the track passes this criterion it is fitted and an

estimate is made of its q/p.

4.4.5 Downstream tracking

The downstream tracking algorithm [55, 56] forms tracks containing hits in the TT (UT) sub-detector and T stations. Each T track is extrapolated back to find the corresponding (x, y) point at the center of the magnet. A track estimate is formed using this point and the nominal interaction point. Hits in the TT consistent with the track estimate are selected. For each TT hit, a new track estimate is formed and consistent x hits are collected. The collection of x hits is fitted in the x-z projection and outliers are removed. Stereo hits are added, the track is refitted and further outliers are removed. Finally, the best track candidate is chosen according to the number of hits it contains and the value of the χ^2 from the fit.

4.4.6 Upstream tracking

The upstream tracking algorithm [57] forms tracks containing hits in the VELO and TT (UT) sub-detectors. These are generally low momentum tracks that will be bent out of acceptance by the magnet. It is executed at the end of the tracking sequence using VELO tracks which have not been upgraded to long tracks by any of the preceeding algorithms.

Each VELO track is linearly extrapolated to the TT. Search windows are opened in each layer and the distance Δx between the track and each hit is calculated. These Δx values are scaled to a reference plane at the center of the TT. A Hough transform is utilised to associate the selected hits to the VELO track. Hits from the same particle are expected to be projected to the same x position in the reference plane while random hits should be uniformly distributed. This procedure is shown schematically in Fig. 18. An explicit assumption is made that every associated hit should lie on the same side of the extrapolated VELO track in the x-z plane.

Each track candidate is fitted with a simplified χ^2 minimisation and the q/p of the track is estimated. Due to the fringe $\vec{B}$ field between the VELO and the TT a momentum estimate of $\delta p/p \sim$ 15% is possible. The best track candidate is chosen based on the number of TT layers containing hits and the χ^2 of the simplified fit. Each of the VeloTT tracks is subsequently fitted with a Kalman filter, described in Sec. 4.6, in order to obtain the most accurate estimates of

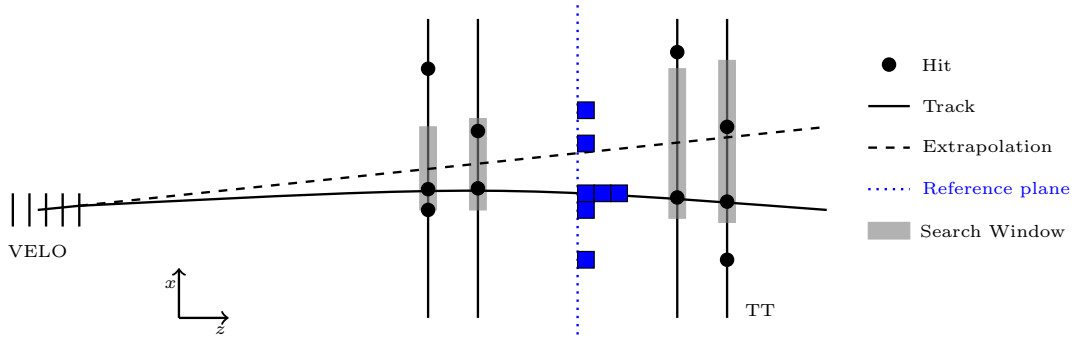


Figure 18: A Hough transform is utilised to associate the selected TT hits to the VELO track. Hits from the same particle are expected to be projected to the same x position in the reference plane while random hits should be uniformly distributed.

track parameters along with their corresponding covariance matrices.

4.5 Clone removal

As there are two independent algorithms to produce long tracks and several track types are subtracks of other types, it is necessary to avoid or remove duplicate tracks found by multiple algorithms. This is accounted for in two different ways. Some algorithms are only allowed to use tracks or hits that have not been previously used. When there is a significant overlap of hits between two tracks, the track with the smaller number of hits is discarded.

4.6 Track fit

The purpose of the track fit is to obtain the most accurate estimates of track parameters along with their corresponding covariances. Track parameters are used to match to particle identification objects (e.g. Cherenkov rings), find primary and secondary vertices and calculate the kinematics and invariant masses of particle combinations. The track χ^2 is used to select good quality tracks.

A Kalman filter is used to fit the tracks. With this approach, multiple scattering is taken into account as process noise and corrections due to energy losses are applied [58]. The transport through the magnetic field is evaluated using a Runge-Kutta method. The propagation and projection functions are linearised around a reference track state using a Taylor expansion [59].

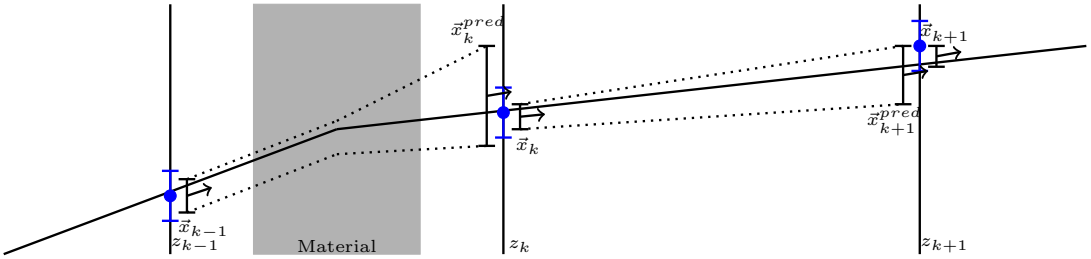


Figure 19: A schematic diagram of the Kalman filter showing the prediction of a state z_k given a state at z_{k-1}. The state is at z_k is subsequently updated with information of the measurement at this position. This process is repeated until all measurements have been added.

Table 3: The simulated samples used in the following sections to benchmark the performance of the upstream tracking algorithms for both the LHCb Upgrade and LHCb Run 2.

	Decay Mode	$\sqrt{s}$ [TeV]	ν	Bunch spacing [ns]
Run 2	$B_s^0 \to \phi\phi$, min. bias	13	1.9	25
Upgrade	$B^0 \to K^{*0}\mu^+\mu^-$, min. bias	14	7.6	25

Track candidates can be considered as a collection of track states (initially provided by the individual tracking algorithms) and measurements (tracking station hits). The Kalman filter can be divided into two steps, shown schematically in Fig. 19. Firstly, the parameters of a state at z_k are predicted given a state at z_{k-1}. Next, the state at z_k is updated with information of the measurement at this position. These two steps are repeated until all the measurements have been added. In LHCb, the filter is run in both the forward and the backward directions and the average is taken for smoothing.

4.7 Tracking performance

The figures of merit used to evaluate the performance of tracking algorithms are the reconstruction efficiency, ghost rate and execution time, which are determined using simulated events. The simulated samples used in the following chapters to benchmark the performance of the upstream tracking algorithms for both the LHCb Upgrade and LHCb Run 2 are shown in Table 3.

4.7.1 Reconstruction efficiency

The reconstruction efficiency is measured by comparing the number of correctly reconstructed tracks with the number of tracks defined to be reconstructible. This is made possible using truth information available in simulated samples. Within the LHCb framework the following definitions are used

- A particle is reconstructible as a VELO track if there are hits associated to it in at least three r and three ϕ VELO sensors.
- A particle is reconstructible as a long track if it fulfils the requirements to be reconstructible in both the VELO sub-detector and the T stations.
- A particle is considered reconstructed if at least 70% of both the VELO and T station hits on a track are associated to it and the track has no more than 1 wrongly associated TT (UT) hit.

The reconstruction efficiency is defined as

$$\epsilon_{rec} = \frac{N_{\text{reconstructible and reconstructed}}}{N_{\text{reconstructible}}}. \tag{4.2}$$

When calculating the efficiency of the VeloTT (VeloUT) algorithm itself, particles are required to be reconstructible as a long track, have been correctly reconstructed in the VELO and have a matched TT (UT) hit in at least 3 TT (UT) layers.

When considering the effect of using VeloTT (VeloUT) tracks as input to the Forward algorithm, no requirement is made that the particle has associated TT (UT) hits or has been correctly reconstructed in the VELO. Therefore, any inefficiency contains contributions from both the VeloTT (VeloUT) algorithm and the acceptance of the TT (UT) detector.

Further requirements are made to both the numerator and the denominator

- The particle is required not to be an electron.
- The pseudorapidity of the particle must lie between 2 and 5.
- The particle is required to be b hadron daughter.
- The particle must have $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}/c$.

These requirements are chosen as b hadrons with a large p_{T} that decay within the LHCb acceptance are of foremost interest within the context of the LHCb software trigger. Electrons are neglected when studying track reconstruction efficiencies, as they are more challenging to reconstruct due to Bremsstrahlung.

4.7.2 Ghost rate and clone tracks

A ghost track is a track that has no matching simulated particle. The ghost rate is defined as

$$\text{Ghost rate} = \frac{N_{\text{ghost tracks}}}{N_{\text{tracks}}}. \tag{4.3}$$

In all cases the following requirements are applied to both the numerator and denominator

- ✦ The pseudorapidity of the track must lie between 2 and 5.
- ✦ Tracks are required to have $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}/c$.

A track matched to a simulated particle with at least one other associated track is said to be a clone.

4.7.3 Execution time

The execution times of the individual algorithms are measured by the LHCb event reconstruction application, Brunel. As the timing is machine dependent, the same machine is used for each measurement to facilitate direct comparisons. A simulated "minimum bias" sample is used in order to not give undue weight to a certain kind of event.

5 Upstream tracking for the LHCb upgrade

5.1 Motivation

The LHCb upgrade will feature a trigger-less readout system allowing the full rate of visible interactions to be processed by a purely software trigger. Such a software trigger allows great flexibility in designing selections and efficient triggering of low momentum tracks normally beyond the scope of a hadron collider. It also places strict requirements on the execution time of the pattern recognition algorithms that run within it.

The existing reconstruction sequence was not able to achieve the required timing performance due to the vast combinatorics present in upgrade conditions. Therefore, a novel idea was proposed to reduce the execution time by using upstream tracks as an intermediate step within the reconstruction sequence [60].

The advantages of using upstream tracks rather than VELO tracks as input to the Forward tracking algorithm arise from the extra information obtained concerning the momentum and charge of the track. Using the momentum information, a preselection on the p_{T} of the track can be performed. Subsequently, for tracks passing the p_{T} requirement, the charge can be used to open smaller search windows downstream of the magnet. This leads to a greatly reduced execution time and ghost rate of the tracking sequence.

In order to achieve the desired improvements within the reconstruction sequence, the upstream algorithm itself must perform with a high reconstruction efficiency, low ghost rate and minimal execution time. Any loss of efficiency will be propagated to the next step of the sequence.

5.2 Initial peformance

The inital version of the VeloUT algorithm for the LHCb upgrade was a replication of the VeloTT algorithm used during Run I, described in Sec. 4.4.6. The aim of this VeloTT algorithm was to reconstruct low momentum tracks that are bent out of acceptance by the magnet. As such, it was executed at the end of the tracking sequence only using VELO tracks which had not been upgraded to long tracks by any of the preceeding algorithms.

The reconstruction efficiency, ghost rate and execution time for the initial version (`v1r2`) of the VeloUT algorithm are shown in Table 4. The reconstruction

Table 4: Reconstruction efficiency, ghost rate and execution time for the initial version (`v1r2`) of the VeloUT algorithm.

VeloUT	Efficiency [%]	Ghost rate [%]	Timing [ms]
v1r2	93.94	7.21	27.20

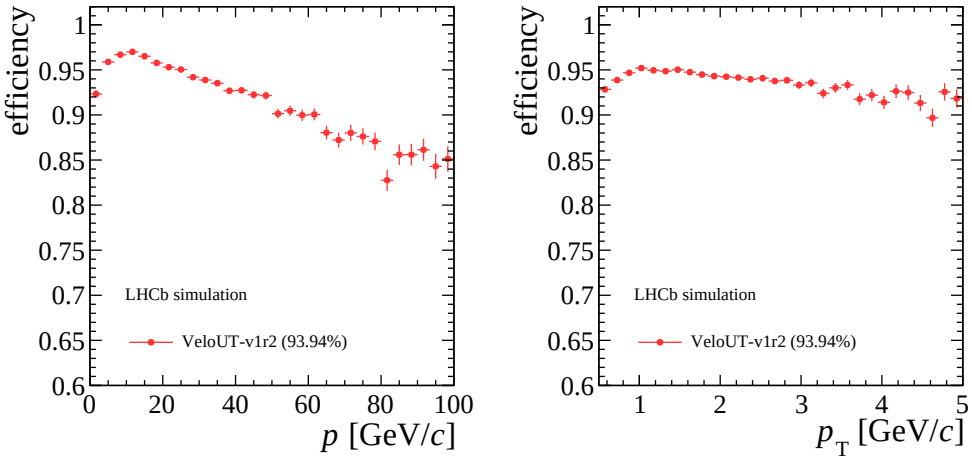


Figure 20: The reconstruction efficiency as a function of p and p_{T} for the initial version (`v1r2`) of the VeloUT algorithm. There is a clear drop in the reconstruction efficiency as a function of p.

efficiency as a function of p and p_{T} are shown in Fig. 20. The ghost rate as a function of p and p_{T} are shown in Fig. 21.

The execution time of 27.20 ms is too slow for the algorithm to be used in the context of a software trigger. For reference, the track reconstruction in the VELO takes $\sim$ 1.8 ms. In order to be included, the VeloUT algorithm should perform with a comparable or reduced execution time.

The reconstruction efficiency is also too low for the algorithm to be used in the context of a software trigger as any inefficiency will be propagated to the next step of the sequence. Furthermore, the efficiency is observed to decrease as a function of p. This is unusual as higher p tracks should bend less in the fringe magnetic field and be simpler to reconstruct.

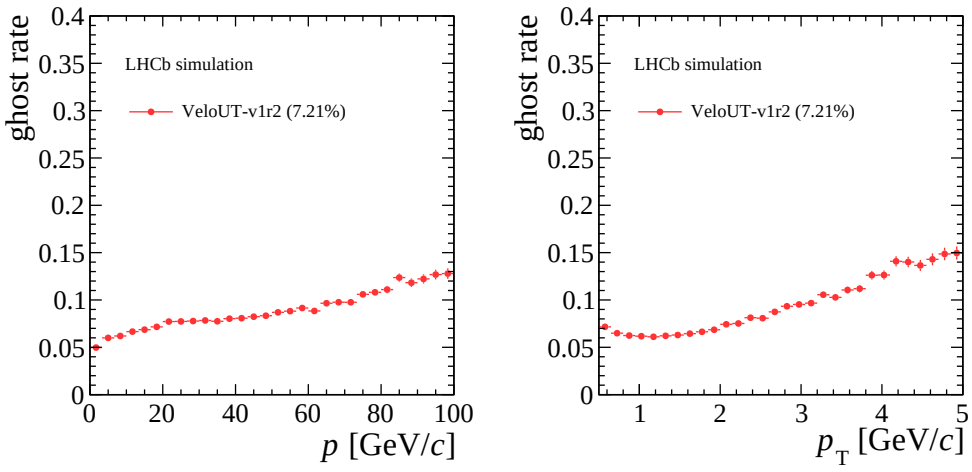


Figure 21: The ghost rate as a function of p and p_{T} for the initial version (`v1r2`) of the VeloUT algorithm.

5.3 Improvements

In order to meet the requirements of the software trigger a number of improvements were made to the VeloUT algorithm. This involved numerous C++ optimisations as well as changes to the logic. The three changes which had the largest effect are described in detail below.

5.3.1 Binary searches

Previously, the UT hits in each layer were sorted into 12 distinct detector regions. This meant that for each layer, each detector region was looped over and if it was compatible with the extrapolated VELO track each hit within that region was looped over. In the new version, the hits in each layer are sorted by their x position at $y = 0$. For each layer, a binary search is performed to find the first hit which is within the search window. The hits are then looped over until that requirement is no longer satisfied. The new method greatly reduces the execution time.

5.3.2 Hit clustering

The inital version of the VeloUT algorithm used a Hough transform based on the distance of the hit from the linear extrapolation of the VELO track to find cluster candidates. It required that all hits were located on one side of the extrapolated

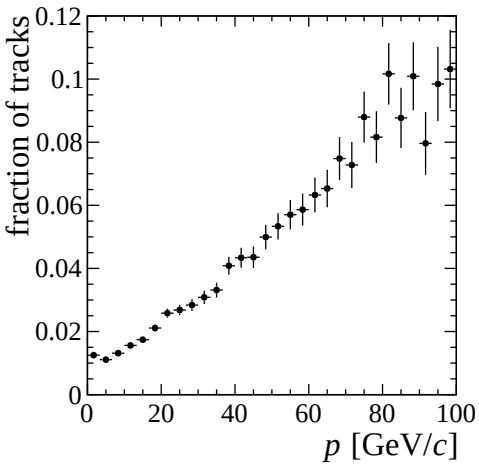


Figure 22: The fraction of truth matched tracks with at least one hit on the opposite side of the linearly extrapolated VELO track in the x-z plane as a function of p.

VELO track in the x-z plane. While this is a good assumption for low p tracks, it is not the case for high p tracks. For high p tracks, a significant number have at least one hit on the opposite side of the linearly extrapolated VELO track, as shown in Fig. 22. This caused the track reconstruction efficiency to fall with increasing p.

A new method for clustering UT hits was developed in order to remedy this flaw and is shown schematically in Fig. 23. In this new method, UT hits are clustered by first forming doublets (two hits in the same station but in different layers), and then extending those doublets to the opposite station and searching for compatible hits to form triplets or quadruplets.

A doublet is formed by taking one hit in the first layer and another in the second layer. The x-slope of the doublet is calculated and if it is below a given threshold the doublet is linearly extrapolated to the third layer where a tolerance window is opened. If there are compatible hits within this window, triplets are formed. For each triplet, the doublet is linearly extrapolated to the fourth layer. A reduced tolerance window is opened and compatible hits are used to form quadruplets. If no quadruplets are formed from any of the triplets, triplets are also searched for with the original doublet and hits in the fourth layer. This process is repeated for every doublet combination.

In order to account for missing hits in the UT detector, if no quadruplets have been formed the clustering sequence is run in reverse starting with a doublet in

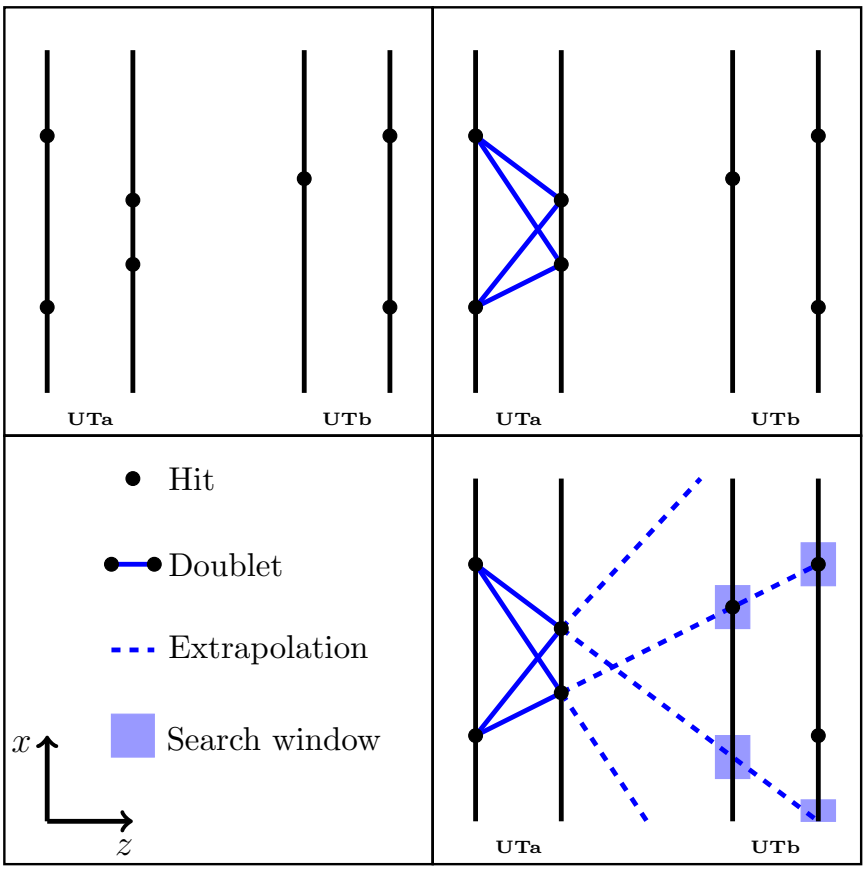


Figure 23: A schematic view of the clustering of UT hit candidates. Doublets in the first two layers are formed and then linearly extrapolated to the third and fourth layers to form triplets and quadruplets.

the third and fourth layers.

The tolerance window in x around the extrapolated x position of the doublet was tuned in simulation. Using simulated particles and their associated UT hits the difference, Δx, between the linearly extrapolated x position of a doublet and the x position of an associated hit in a given layer can be found. The distributions used for the tuning are shown in Fig. 24.

5.3.3 Track fit

The initial version of the VeloUT algorithm fitted each of the VeloUT tracks with a Kalman filter, described in Sec. 4.6, in order to obtain the most accurate estimates of track parameters along with their corresponding covariances. This was very costly in terms of execution time and did not provided any significant

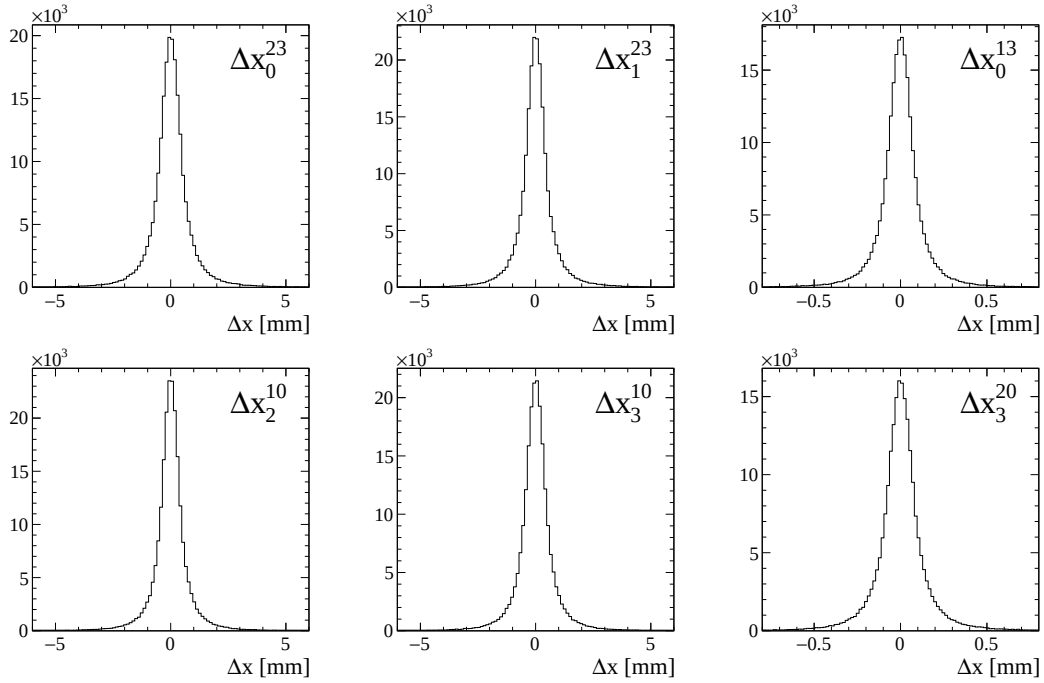


Figure 24: The difference Δx_c^{ab} between the linearly extrapolated x position of a doublet and the x position of an associated hit in a given layer where a and b denote the two layers from which the slope has been calculated and c denotes the layer to which the extrapolation is being performed.

improvement to the momentum or charge estimation. The Kalman filter was removed and the momentum and charge information taken from the simplified fit described in Sec. 4.4.6, leading to a vast improvement in the execution time.

5.4 Upgrading to long tracks

VeloUT tracks rather than VELO tracks will be used as input to the Forward tracking algorithm in the LHCb Upgrade. Using the charge and momentum information of the VeloUT track it is possible to make smarter selections on the input tracks and T-station hits considered by the Forward algorithm [49, 60].

A preselection of $p_{\mathrm{T}} > 400\,\mathrm{MeV}/c$ reduces the number of input tracks passed to the Forward tracking by a factor three compared to using VELO tracks. The charge information allows smaller, asymmetric search windows to be opened in the T-stations, reducing the hit multiplicity by a factor two. A small window is also opened on the 'wrong' side of the linear extrapolation for high p_{T} track candidates as they are more likely to have been assigned the incorrect charge. The optimised

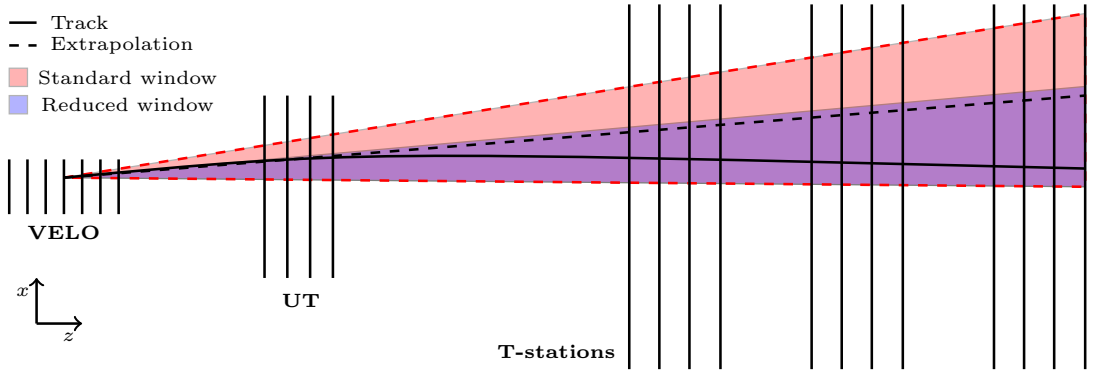


Figure 25: The search windows opened by the Forward algorithm with and without the charge and momentum information of the VeloUT candidates. The charge information allows smaller, asymmetric search windows to be opened. A small window is also opened on the 'wrong' side of the linear extrapolation for high p_{T} candidates as they are more likely to have been assigned the incorrect charge.

search windows are shown schematically in Fig. 25. These two advancements lead to a greatly reduced execution time and ghost rate of the Forward algorithm. In order to prevent a loss in efficiency due to the central acceptance of the UT, any VELO tracks that are linearly extrapolated within the central hole are passed directly to the Forward tracking algorithm.

5.5 Performance

5.5.1 VeloUT

The track reconstruction efficiency, ghost rate and execution time of the VeloUT algorithm for both the initial version (`v1r2`) and the optimised version (`v2r2`) are shown in Table 5. The track reconstruction efficiency as a function of p and p_{T} are shown in Fig. 26. The ghost rate as a function of p and p_{T} are shown in Fig. 27. The optimised version shows large improvements in terms of track reconstruction efficiency and execution time. The increase in reconstruction efficiency is most evident at high p where the initial version shows a negative trend for increasing p. There is also a slight increase in the ghost rate. However, this is of lesser importance as the ghost rate can be further reduced during offline analysis.

Table 5: The performances of both the initial version (`v1r2`) and the optimised version (`v2r2`) of the VeloUT algorithm in terms of track reconstruction efficiency, ghost rate and execution time.

VeloUT	Efficiency [%]	Ghost rate [%]	Timing [ms]
v1r2	93.94	7.21	27.20
v2r2	98.69	8.00	0.81

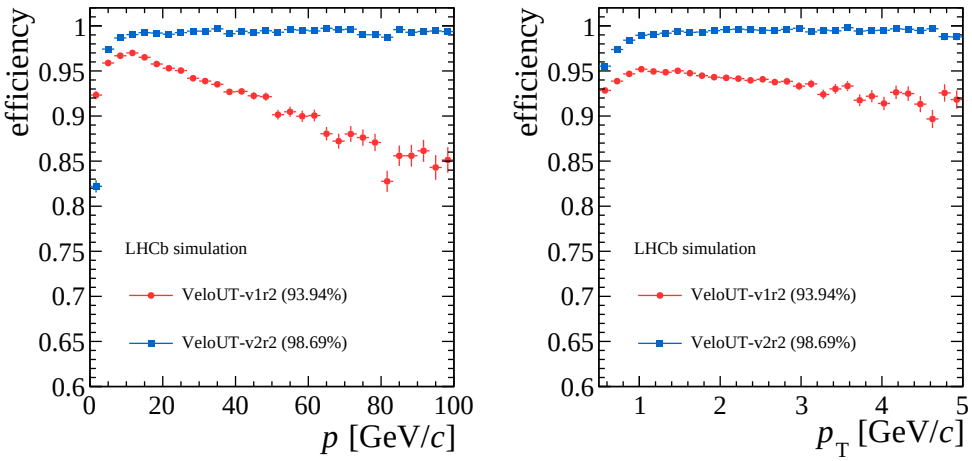


Figure 26: The reconstruction efficiency as a function of p and p_{T} for both the initial version (`v1r2`) and the optimised version (`v2r2`) of the VeloUT algorithm.

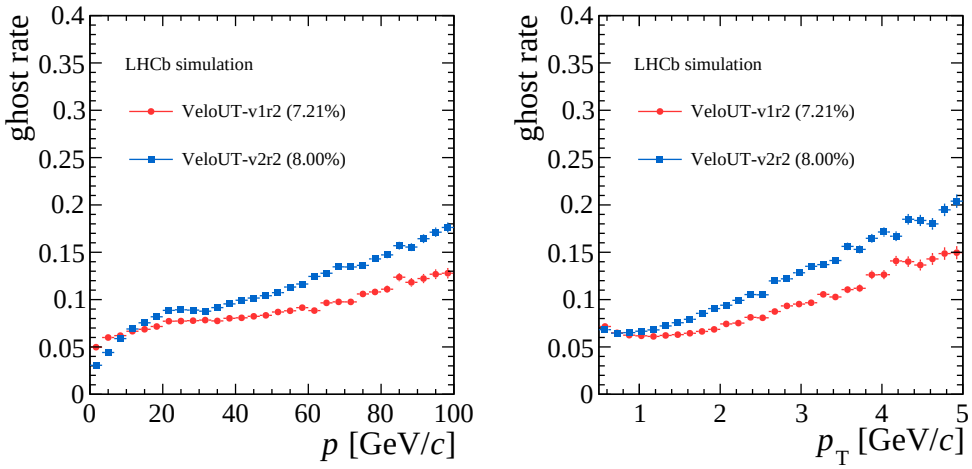


Figure 27: The ghost rate as a function of p and p_{T} for both the initial version (`v1r2`) and the optimised version (`v2r2`) of the VeloUT algorithm.

Table 6: The performances of the Forward algorithm using VELO or VeloUT tracks as input in terms of track reconstruction efficiency, ghost rate and execution time.

	Efficiency [%]	Ghost rate [%]	VeloUT [ms]	Forward [ms]
Velo-Forward	94.10	41.55	-	18.28
VeloUT-Forward	93.37	14.08	0.81	3.45

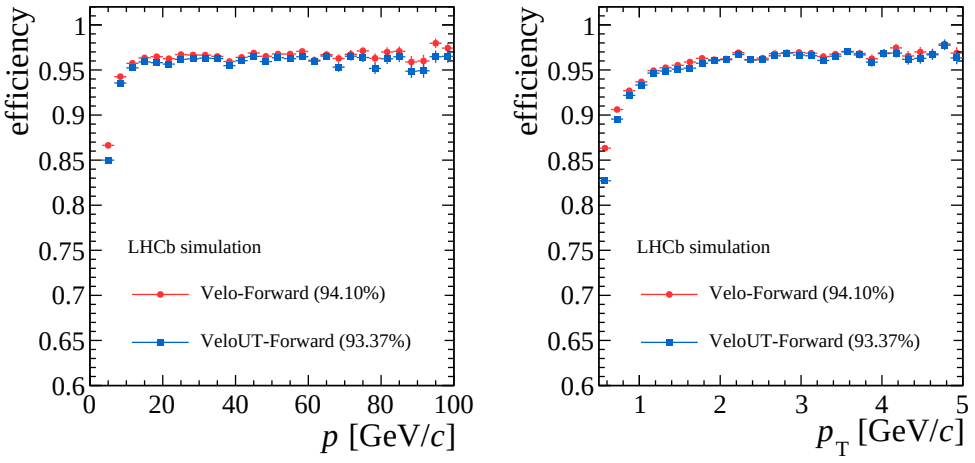


Figure 28: The track reconstruction efficiency of the Forward algorithm using VELO or VeloUT tracks as a function of p and p_{T}.

5.5.2 Forward

The track reconstruction efficiency, ghost rate and execution time of the Forward algorithm using VELO or VeloUT tracks as input are shown in Table 6. The track reconstruction efficiency as a function of p and p_{T} are shown in Fig. 28. The ghost rate as a function of p and p_{T} are shown in Fig. 29. The use of VeloUT tracks as input to the Forward algorithm drastically reduces the ghost rate and execution time. This comes at a small cost in the track reconstruction efficiency. This efficiency can be recovered by the Matching algorithm

5.6 Summary

The majority of LHCb analyses use long tracks, which traverse the full tracking detector. These tracks will be reconstructed by the Forward tracking algorithm at the software trigger level in the LHCb upgrade. Previously, VELO tracks were used as input to the Forward algorithm. However, a novel method was investigated

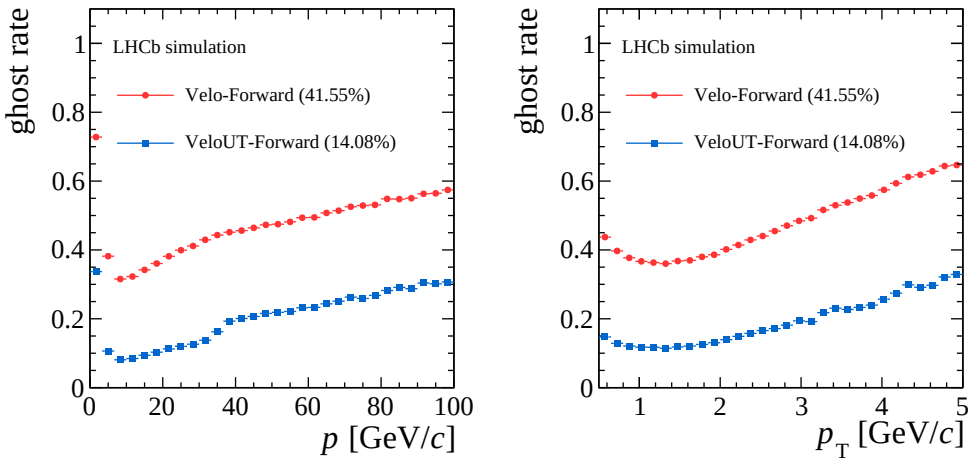


Figure 29: The ghost rate of the Forward algorithm using VELO or VeloUT tracks as a function of p and p_{T}.

to use upstream tracks, reconstructed by the VeloUT algorithm, as input. The vast improvements in the performance of the VeloUT algorithm and the subsequent improvement to the overall tracking sequence led to the algorithm being adopted into the default tracking sequence for the LHCb Upgrade [47, 49]. It will play an important role in allowing LHCb to become the first hadron collider experiment to operate a software-only trigger at the full event rate.

6 Upstream tracking for LHCb Run 2

6.1 Motivation

Following the improved performance achieved using VeloUT tracks within the tracking sequence of the LHCb Upgrade, outlined in Chapter 5, a similar strategy was proposed for LHCb Run 2. This strategy involved using VeloTT tracks as input for the Forward tracking at the first stage of the software trigger. A new VeloTT algorithm for Run 2 [61] was created based on the optimised VeloUT algorithm, taking into account the slight differences in geometry between the two sub-detectors. In the following chapter, the improved performance of the VeloTT algorithm for Run 2 compared to Run 1 will be shown along with the improved performance achieved when using VeloTT tracks rather than VELO tracks as input to the Forward tracking algorithm.

6.2 Performance

6.2.1 VeloTT

The track reconstruction efficiency, ghost rate and execution time of the VeloTT algorithm for Run 1 and Run 2 are shown in Table 7. The track reconstruction efficiency as a function of p and p_{T} are shown in Fig. 30. The ghost rate as a function of p and p_{T} are shown in Fig. 31. The Run 2 implementation shows large improvements in terms of track reconstruction efficiency and execution time. The increased track reconstruction efficiency is most evident at high p as the Run 1 implementation shows a negative trend for increasing p. There is also a slight increase in the ghost rate. However, this is of lesser importance as the ghost rate can be further reduced during offline analysis.

Table 7: The performances of the VeloTT algorithms for Run 1 and Run 2 in terms of track reconstruction efficiency, ghost rate and execution time.

VeloTT	Efficiency [%]	Ghost rate [%]	Timing [ms]
Run 1	92.74	7.21	32.50
Run 2	97.77	11.60	0.50

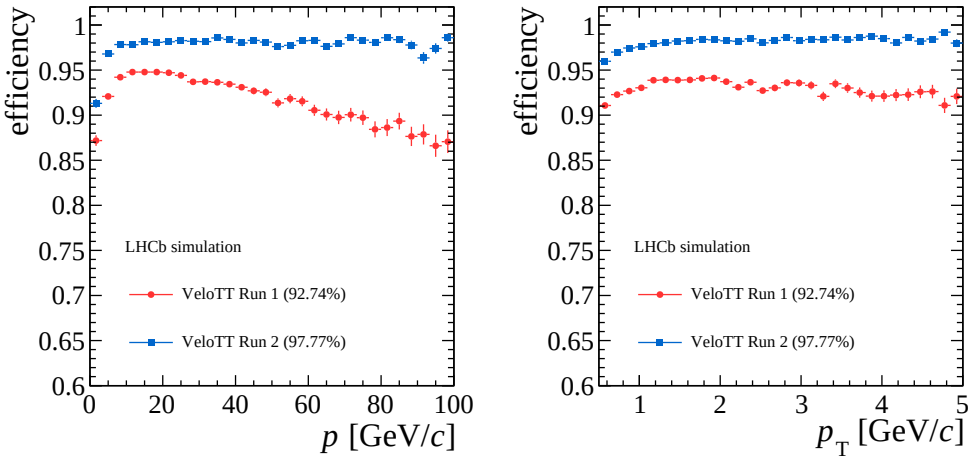


Figure 30: The track reconstruction efficiency of the VeloTT algorithms for Run 1 and Run 2 as a function of p and p_{T}.

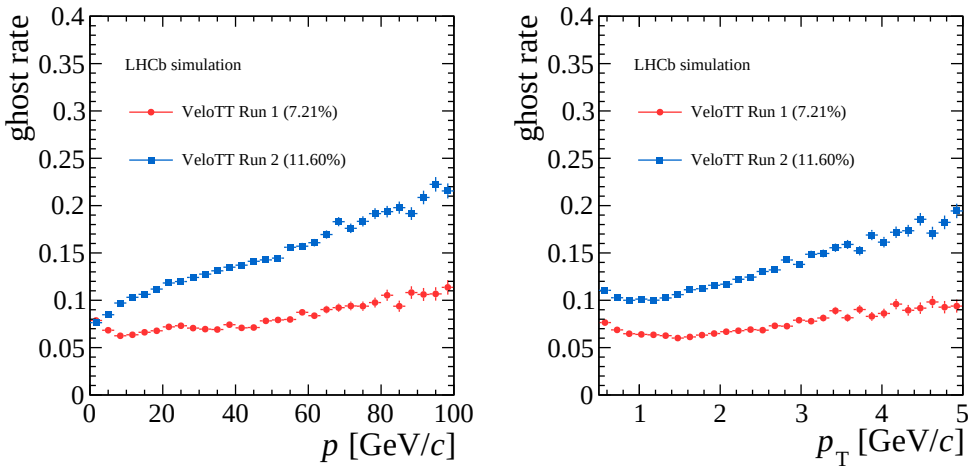


Figure 31: The ghost rate of the VeloTT algorithms for Run 1 and Run 2 as a function of p and p_{T}.

6.2.2 Forward

The track reconstruction efficiency, ghost rate and execution time of the Forward algorithm taking VELO or VeloTT tracks as input are shown in Table 8. The track reconstruction efficiency as a function of p and p_{T} are shown in Fig. 32. The ghost rate as a function of p and p_{T} are shown in Fig. 33. The use of VeloTT tracks as input drastically reduces the ghost rate and execution time of the Forward algorithm. This comes at a small cost in the track reconstruction efficiency. This

Table 8: The performances of the Forward algorithm using VELO or VeloTT tracks as input in terms of track reconstruction efficiency, ghost rate and execution time.

	Efficiency [%]	Ghost rate [%]	VeloTT [ms]	Forward [ms]
Velo-Forward	93.15	46.86	-	13.71
VeloTT-Forward	89.23	17.13	0.50	4.08

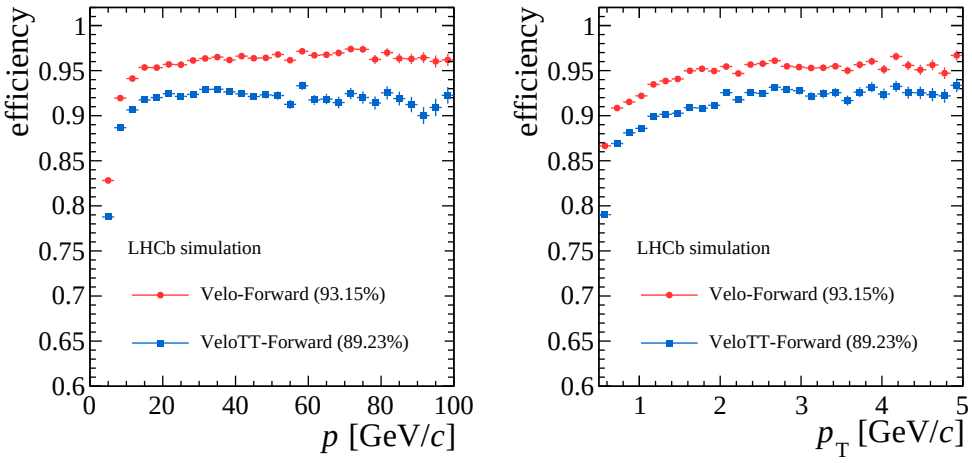


Figure 32: The track reconstruction efficiency of the Forward algorithm using VELO or VeloTT tracks as a function of p and p_{T}.

efficiency is retrieved at the second stage of the software trigger [40].

6.3 Summary

The enhanced performance achieved when using VeloTT tracks as input to the Forward tracking led to the algorithm being included in the reconstruction sequence at the first stage of the software trigger for Run 2. The reduction in the execution time by a factor three makes it possible to remove any IP requirement on VELO tracks and to loosen the p_{T} threshold of the Forward tracking from 1.2 GeV/c to 0.5 GeV/c in the first stage of the software trigger. This greatly improves the signal efficiency for charm physics and allows lifetime unbiased triggers for hadronic final states for the first time [40]. It also plays an important role in allowing physics analyses to be performed directly on the output of the software trigger, as described in Sec 3.3.

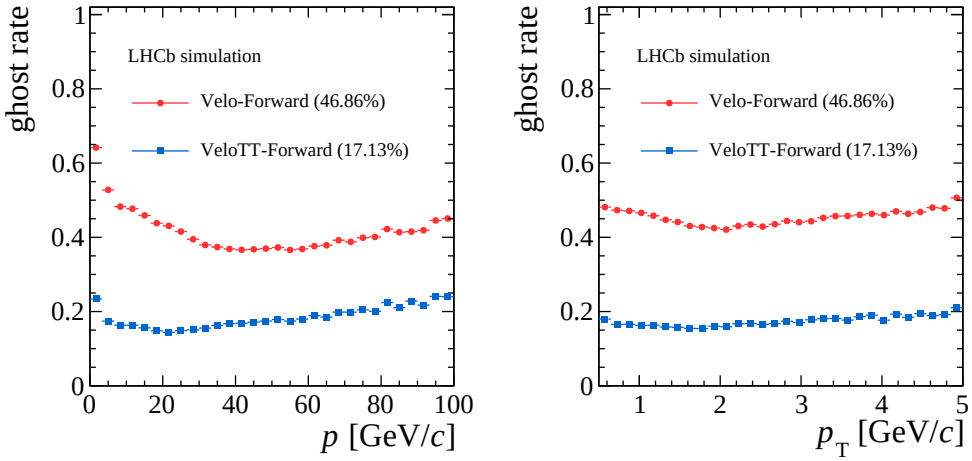


Figure 33: The ghost rate of the Forward algorithm using Velo or VeloTT tracks as a function of p and p_{T}.

7 Study of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region

7.1 Introduction

The decay $B^0 \to K^+\pi^-\mu^+\mu^-$ is a flavour-changing neutral-current process that proceeds via a $b \to s\mu^+\mu^-$ transition. In the SM, the leading order transition amplitudes are described by electroweak penguin or box diagrams. In extensions to the SM, new heavy particles can contribute to loop diagrams and modify observables such as branching fractions and angular distributions. In this chapter, the study of $B^0 \to K^+\pi^-\mu^+\mu^-$ decays in the $K^*_{0,2}(1430)^0$ region, published in Ref. [62], is presented.

7.1.1 Previous $b \to s\mu^+\mu^-$ measurements

The previous angular analyses of $B^0 \to K^+\pi^-\mu^+\mu^-$ performed by the LHCb collaboration [63, 64, 65, 66] focused on the $K^+\pi^-$ invariant mass range $796 < m(K^+\pi^-) < 996\,\text{MeV}/c^2$ where the decay proceeds predominantly via the P-wave process $K^*(892)^0 \to K^+\pi^-$. A global analysis of the CP-averaged angular observables measured in the LHCb Run 1 data sample indicated differences with SM predictions at the level of 3.4 standard deviations [66]. The results of the measurement of the observable P_5', which exhibits a local deviation from the SM predictions, is shown in Fig. 34.

This set of measurements is part of a pattern of discrepancies with respect to SM predictions that have been observed in $b \to s\mu^+\mu^-$ transitions. For example, the measured differential branching fractions of the decays $B^0_s \to \phi\mu^+\mu^-$ [68], $\Lambda^0_b \to \Lambda\mu^+\mu^-$ [69] and $B^+ \to K^+\mu^+\mu^-$ [70] all lie below their corresponding SM predictions. Furthermore, the ratio $R_K = \mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+e^+e^-)$, which is a test of lepton flavour universality, was also measured to be 2.6 standard deviations from its SM prediction of unity [71].

This pattern of discrepancies can be intepreted by performing global model-independent fits to $b \to s\mu^+\mu^-$ measurements [72].[5] In the framework of OPE, described in Sec 2.4, these fits can be used to constrain the values of the Wilson coefficients $\mathcal{C}_7$, $\mathcal{C}_9$ and $\mathcal{C}_{10}$. A χ^2 function which quantifies the compatibility of

[5]The global fit in Ref. [72] takes into account 88 measurements of 76 observables by the ATLAS, BaBar, Belle, CDF, CMS and LHCb experiments.

the model with the data, for a given set of values of the Wilson coefficients, is minimised in different scenarios. The NP dependencies are encoded as NP contributions to the Wilson coefficients, $\mathcal{C}_i^{\rm NP} = \mathcal{C}_i - \mathcal{C}_i^{\rm SM}$. The best fit is obtained when allowing NP in $\mathcal{C}_9$ only, yielding a value of $\mathcal{C}_9^{\rm NP} = -1.07$ which correponds to a pull of 3.7 standard deviations from the SM. Figure 35 shows the result of the fit when allowing for NP effects in both $\mathcal{C}_9$ and $\mathcal{C}_{10}$. These results are in good agreement with Ref. [73], which also finds that a negative contribution to $\mathcal{C}_9$ plays a central role in explaining the observed discrepancies.

Many NP models have been proposed to explain this observed tension from the SM in $\mathcal{C}_9$. Such models contain new interactions mediated by a Z' boson [12, 13, 14] or leptoquarks [74, 75, 76]. These interactions can also introduce a violation of lepton flavour universality.

However, it has also been suggested that the contribution of so-called charm-loop effects could be responsible for the observed deviations [77]. As these hadronic effects are mediated via virtual photon exchange, leading to a vector-like coupling to leptons, it is possible they could mimic NP effects in $\mathcal{C}_9$.

Since short distance effects are universal and should appear coherently in all $b \to s\mu^+\mu^-$ transitions, measuring other $b \to s\mu^+\mu^-$ transitions can help to shed light on this situation. To this end, the current chapter describes the study of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in a previously unexplored region of $m(K^+\pi^-)$.

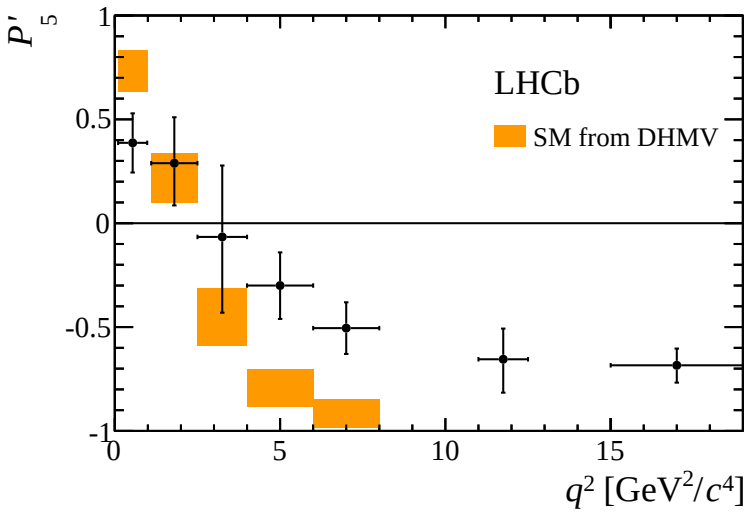


Figure 34: Results of the measurement of the observable P_5' by the LHCb collaboration. The SM predictions are taken from Ref. [67].

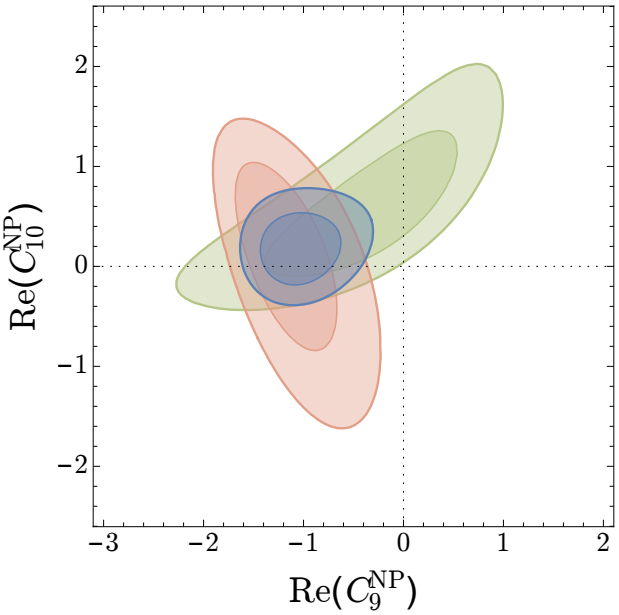


Figure 35: Allowed region in the $\mathrm{Re}(\mathcal{C}_9^{\mathrm{NP}})$–$\mathrm{Re}(\mathcal{C}_{10}^{\mathrm{NP}})$ plane. The blue contours correspond to the 1σ and 2σ best fit regions from the global fit. The red and green contours represent the 1σ and 2σ regions if only the $B^0 \to K^*(892)^0\mu^+\mu^-$ angular observables or only the differential branching fraction measurements are taken into account, respectively. Taken from Ref. [72].

7.1.2 Analysis overview

As the dominant structures in the $K^+\pi^-$ invariant mass spectrum of $B^0 \to K^+\pi^-\mu^+\mu^-$ above the P-wave $K^*(892)^0$ are resonances in the 1430 MeV/c^2 region, this is a natural region to study. The relevant K^{*0} states above the $K^*(892)^0$ mass range are listed in Table 9. Throughout this thesis, the symbol K^{*0} denotes any neutral strange meson in an excited state that decays to a $K^+\pi^-$ final state. In the 1430 MeV/c^2 region, contributions are expected from the S-wave $K_0^*(1430)^0$, P-wave $K^*(1410)^0$ and D-wave $K_2^*(1430)^0$ states, as well as the broad P-wave $K^*(1680)^0$ state.

The $m(K^+\pi^-)$ distribution for $B^0 \to K^+\pi^-\mu^+\mu^-$ decays in the range $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ and $630 < m(K^+\pi^-) < 1630\,\mathrm{MeV}/c^2$ is shown in Fig. 36, where $q^2 \equiv m^2(\mu^+\mu^-)$. The candidates are obtained using the selection described in Sec. 7.3 and the background component is subtracted using the *sPlot* technique [79]. The main structures are observed around the mass of the $K^*(892)^0$ resonance and in the 1430 MeV/c^2 region.

This chapter describes the first measurements of the differential branching frac-

Table 9: Expected resonant contributions above the $K^*(892)^0$ mass range. For each, the spin-parity, J^P, and branching fraction to $K\pi$, $\mathcal{B}(K\pi)$, are given. Taken from Ref. [78].

Resonance	J^P	Mass [MeV/c^2]	Full width [MeV/c^2]	$\mathcal{B}(K\pi)$ [%]
$K^*(1410)^0$	1^-	1414 ± 15	232 ± 21	6.6 ± 1.3
$K_0^*(1430)^0$	0^+	1425 ± 50	270 ± 80	93 ± 10
$K_2^*(1430)^0$	2^+	1432.4 ± 1.3	109 ± 5	49.9 ± 1.2
$K^*(1680)^0$	1^-	1717 ± 27	322 ± 110	38.7 ± 2.5
$K_3^*(1780)^0$	3^-	1776 ± 7	159 ± 21	18.8 ± 1.0
$K_4^*(2045)^0$	4^+	2045 ± 9	198 ± 30	9.9 ± 1.2

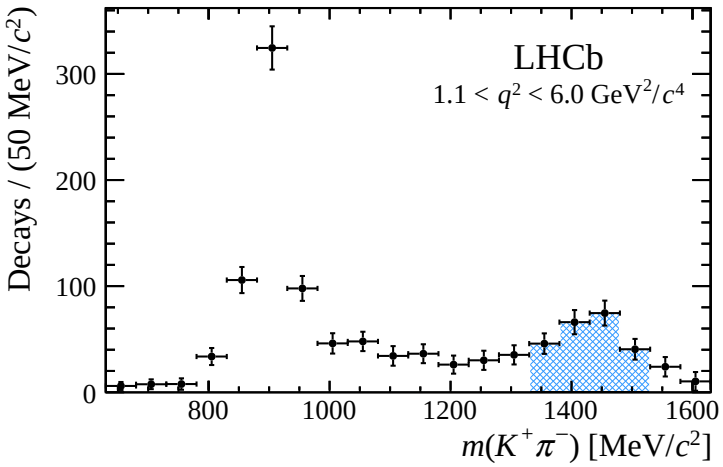
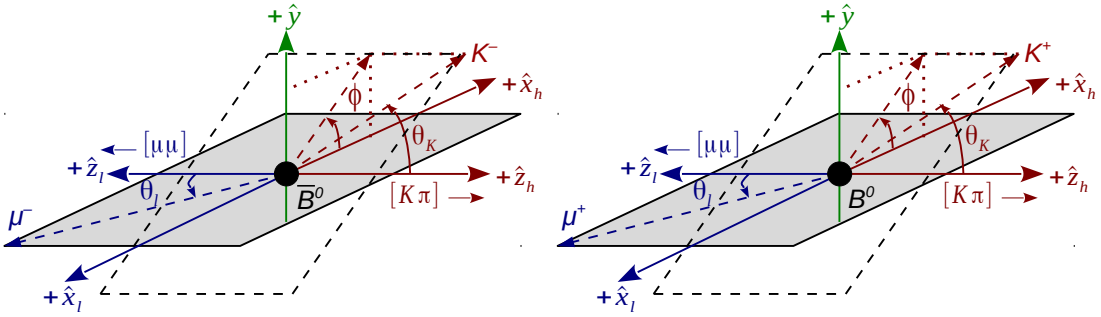


Figure 36: Background-subtracted $m(K^+\pi^-)$ distribution for $B^0 \to K^+\pi^-\mu^+\mu^-$ decays in the range $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$. The region $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 is indicated by the blue hatched area.

tion and angular moments of $B^0 \to K^+\pi^-\mu^+\mu^-$ in the region $1330 < m(K^+\pi^-) < 1530\,\text{MeV}/c^2$. The values of the differential branching fraction are reported in five bins of q^2 between 0.1 and $8.0\,\text{GeV}^2/c^4$, and in the range $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$ for which the angular moments are also measured. The measurements are based on samples of pp collisions collected by the LHCb experiment in Run 1, corresponding to integrated luminosities of $1.0\,\text{fb}^{-1}$ at a centre-of-mass energy of 7 TeV and $2.0\,\text{fb}^{-1}$ at 8 TeV.

7.2 Angular distribution and observables

The final state of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ is fully described by five kinematic variables: three decay angles (θ_ℓ, θ_K, ϕ), $m(K^+\pi^-)$, and q^2. Figure 37a shows



(a) (b)

Figure 37: Angle conventions for the (a) $\overline{B}^0 \to K^-\pi^+\mu^-\mu^+$ (b) $B^0 \to K^+\pi^-\mu^+\mu^-$ as described in Ref. [80]. The leptonic and hadronic frames are back-to-back with a common $\hat{y}$ axis. For the dihedral angle ϕ between the leptonic and hadronic decay planes, there is an additional sign flip $\phi \to -\phi$ compared to previous LHCb analyses [63, 64, 65, 66].

the angle conventions for the $\overline{B}^0$ decay (containing a b quark): the back-to-back leptonic and hadronic systems share a common $\hat{y}$ axis and have opposite $\hat{x}$ and $\hat{z}$ axes. The negatively charged lepton is used to define the leptonic helicity angle θ_ℓ for the $\overline{B}^0$. The quadrant of the dihedral angle ϕ between the dimuon and the $\overline{K}^{*0} \to K^-\pi^+$ decay planes is determined by requiring the azimuthal angle of the μ^- to be zero in the leptonic helicity frame. The azimuthal angle of the K^- in the hadronic helicity frame is then equal to ϕ. Compared to the dihedral angle used in the previous LHCb analyses [63, 64, 65, 66], there is a sign flip, $\phi \to -\phi$, in the convention used here. For the B^0 decay (containing a $\overline{b}$ quark), the charge conjugation is performed explicitly, and the angles are shown in Fig. 37a, where for the B^0, the μ^+ and K^+ directions are used to define the angles. An additional minus sign is added to the dihedral angle when performing the CP conjugation, in order to keep the measured angular observables the same between $\overline{B}^0$ and B^0 in the absence of direct CP violation.

In the limit where q^2 is large compared to the square of the muon mass, the CP-averaged differential decay rate of $B^0 \to K^+\pi^-\mu^+\mu^-$ with the $K^+\pi^-$ system in a S-, P-, or D-wave configuration can be expanded in an orthonormal basis of angular functions $f_i(\Omega)$ as

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2\,\mathrm{d}\Omega} \propto \sum_{i=1}^{41} f_i(\Omega)\Gamma_i(q^2) \quad \text{with} \quad \Gamma_i(q^2) = \Gamma_i^L(q^2) + \eta_i^{L\to R}\,\Gamma_i^R(q^2), \tag{7.1}$$

where $\mathrm{d}\Omega = \mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi$, and L and R denote the (left- and right-handed) chirality of the lepton system [80]. The sign $\eta_i^{L\to R} = \pm 1$ depends on whether f_i changes sign under $\theta_\ell \to \pi + \theta_\ell$. The orthonormal angular basis is constructed out of spherical harmonics, $Y_l^m \equiv Y_l^m(\theta_\ell, \phi)$, and reduced spherical harmonics, $P_l^m \equiv \sqrt{2\pi} Y_l^m(\theta_K, 0)$.

The transversity-basis moments of the 41 orthonormal angular functions are given in Appendix A. The convention is that the amplitudes correspond to the $\overline{B}^0$ decay. The S-, P- and D-wave transversity amplitudes are denoted as $S^{\{L,R\}}$, $H^{\{L,R\}}_{\{0,\parallel,\perp\}}$ and $D^{\{L,R\}}_{\{0,\parallel,\perp\}}$, respectively.

The measured angular observables are averaged over the range $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 and $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$. This q^2 range is part of the large-recoil regime where the recoiling K^{*0} has a relatively large energy, $E_{K^{*0}}$, as measured in the rest frame of the parent B meson. In the limit $\Lambda_{\mathrm{QCD}}/E_{K^{*0}} \to 0$, the uncertainties arising from hadronic effects in the relevant form-factors are reduced at leading order, resulting in more reliable theory predictions [81]. The high-q^2 region above the $\psi(2S)$ resonance is polluted by broad charmonium resonances and is also phase-space suppressed for higher $m(K^+\pi^-)$ masses. Therefore, that region is not considered in this study.

In the present analysis, the first moment, $\Gamma_1(q^2)$, corresponds to the total decay rate. From this, 40 normalised moments for $i \in \{2, ..., 41\}$ are defined as

$$\overline{\Gamma}_i(q^2) = \frac{\Gamma_i(q^2)}{\Gamma_1(q^2)}. \tag{7.2}$$

These form the set of observables that are measured in the angular moments analysis described in Sec. 7.8.

7.3 Candidate selection

The selection of $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates comprises several steps. Firstly, candidates are required to have 'fired' at least one of the specified trigger lines at each of the three stages of the trigger. Subsequently, candidates must pass two sets of requirements: the first is a common selection known as 'stripping' and the second is a loose preselection. Next, combinatorial background candidates are reduced using a multivariate classifier. Finally, exclusive backgrounds are removed with specific vetoes.

7.3.1 Data samples

The Run 1 data sample collected by the LHCb experiment is used for this analysis, corresponding to an integrated luminosity of $3.0\,\text{fb}^{-1}$. The data were recorded in pp collisions at centre-of-mass energies of 7 and 8 TeV during 2011 and 2012, respectively. In addition, a number of simulated samples are used to evaluate possible peaking background contributions and to determine the acceptance correction.

7.3.2 Trigger requirements

At the first stage of the trigger, L0, the event must have been triggered by a single muon from the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay. At the second level of the trigger, HLT1, at least one of two possible trigger lines must have been triggered by a single daughter particle from the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay. At the final level of the trigger, HLT2, at least one of several trigger lines must have been triggered by the daughters of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay: these include both topological and muon triggers.

7.3.3 Stripping and preselection

The stripping is a set of common, cut-based requirements used to loosely select candidates of interest for similar analyses. The stripping line used in this analysis selects candidates of the form $B \to X\mu^+\mu^-$, where X is one or more hadrons. The full set of stripping requirements are shown in Table 10. The boolean variable `IsMuon` is used to select muons and requires the particle to have left hits in a given number of muon stations depending on its measured momentum. A global event cut (GEC) is applied on the number of hits in the SPD to reject very high multiplicity events.

A loose preselection of candidates is performed to remove pathological events. The full set of preselection requirements are shown in Table 11. The boolean variable `hasRich` requires the particle to have information from the RICH detectors. The angles θ and θ_{pair} represent the opening angle of a track from the beam and the opening angle between a track pair, respectively. The variables $\langle X \rangle$, $\langle Y \rangle$ and $\langle Z \rangle$ denote the mean primary vertex position.

Table 10: Stripping requirements applied to $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates.

	Requirement
B	$4600 < M < 7000\,\text{MeV}/c^2$ vertex quality $\chi^2/\text{ndf} < 6$ vertex separation $\chi^2 > 121$ IP $\chi^2 < 16$ DIRA $< 14\,\text{mrad}$
K^{*0}	$M < 6200\,\text{MeV}/c^2$ vertex quality $\chi^2/\text{ndf} < 12$ vertex separation $\chi^2 > 9$
$\mu^+\mu^-$	$M < 7100\,\text{MeV}/c^2$ vertex quality $\chi^2/\text{ndf} < 12$
μ	`IsMuon` $\text{DLL}_{\mu\pi} > -3$
tracks	Ghost probability < 0.4 IP $\chi^2 > 9$
GEC	SPD multiplicity < 600

Table 11: Preselection requirements applied to $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates.

	Requirement
K	`hasRich` $\text{DLL}_{K\pi} > -5$
π	`hasRich` $\text{DLL}_{K\pi} < 25$
track track pairs	$0 < \theta < 400\,\text{mrad}$ $\theta_{\text{pair}} > 1\,\text{mrad}$
PV	$\lvert X - \langle X\rangle\rvert < 5\,\text{mm}$ $\lvert Y - \langle Y\rangle\rvert < 5\,\text{mm}$ $\lvert Z - \langle Z\rangle\rvert < 200\,\text{mm}$

7.3.4 Multivariate classifier

A multivariate classifier is used to reduce the level of combinatorial background. A Boosted Decision Tree (BDT) classifier [82], with the Adaboost algorithm [83] is employed. The BDT was originally developed for the angular analysis of $B^0 \to K^*(892)^0\mu^+\mu^-$ [66].

The BDT uses a combination of kinematic and PID input variables: the B^0 candidate lifetime, the B^0 p and p_{T}, the B^0 DIRA, the B^0 vertex quality χ^2, the $\text{DLL}_{K\pi}$ of the kaon and pion, the $\text{DLL}_{\mu\pi}$ of the muons and the isolation of the four final state particles. The isolation exploits the idea that the daughters of 'true'

$B^0 \to K^+\pi^-\mu^+\mu^-$ candidates should better isolated from other tracks in the event than those from combinatorial candidates, which originate from two distinct B mesons.

The training uses $B^0 \to J/\psi\, K^+\pi^-$ candidates as a proxy for the signal and $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates from the upper mass sideband as a proxy for the background. A k-folding approach [84] is employed to allow the full dataset to be used for testing and training in an unbiased way. The dataset is randomly divided into $k = 10$ samples of equal size. Ten classifiers are trained, each using nine signal and nine background samples. Each classifier is then applied to the remaining sample that was omitted during the training. This approach allows 90% of the dataset to be used when training each classifier, compared to only 50% when taking the traditional approach of dividing the dataset in two.

7.3.5 Exclusive backgrounds

Several additional requirements are applied to remove contributions from decay modes that will peak at, or near to, the signal peak and therefore distort the distributions of $\cos\theta_\ell$, $\cos\theta_K$ and ϕ. These decay modes include $\Lambda_b^0 \to pK^-\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$ as well as misidentified $B^0 \to J/\psi\, K^+\pi^-$, $B^0 \to \psi(2S)K^+\pi^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$. The full set of requirements are presented in Table 12. These requirements reduce the expected contamination from exclusive background candidates to the level of 2% of the signal yield,

A background from $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decays arises when the p is reconstructed as either of the hadron candidates. Candidates are rejected using PID information and alternative mass hypotheses. For the case when the p is reconstructed as the π, a new mass is constructed where the π is given the p mass hypothesis. Likewise for the case when the p is reconstructed as the K, a new mass is constructed where the K is given the p mass hypothesis, and the π the K mass hypothesis. These new mass hypotheses are denoted as $m_{(\pi\to p)K\mu\mu}$ and $m_{(K\to p)(\pi\to K)\mu\mu}$ repectively. Candidates from $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decays are expected have $m_{(\pi\to p)K\mu\mu}$ or $m_{(K\to p)(\pi\to K)\mu\mu}$ consistent with the known Λ_b^0 mass.

A further peaking background contribution can be formed if a π^- from elsewhere in the event is added to a genuine $B^+ \to K^+\mu^+\mu^-$ decay to form a four-track final state. As $B^+ \to K^+\mu^+\mu^-$ candidates will accumulate at the nominal B^+ mass, these candidates are expected to reside in the upper $m(K^+\pi^-\mu^+\mu^-)$ side-

Table 12: Requirements applied to veto exclusive backgrounds. All invariant masses have the units MeV/c^2.

Decay mode	Mis-id	Veto
$\Lambda_b^0 \to pK^-\mu^+\mu^-$	$p \to \pi$	$5575 < m_{(\pi\to p)K\mu\mu} < 5665$ and $\pi\text{DLL}_{p\pi} > 0$
	$p, K \to K, \pi$	$5575 < m_{(K\to p)(\pi\to K)\mu\mu} < 5665$ and $\pi\text{DLL}_{K\pi} > 0$
$B^+ \to K^+\mu^+\mu^-$	Random π	$m_{K\pi\mu\mu} > 5380$ and $5220 < m_{K\mu\mu} < 5340$
$B^0 \to J/\psi\, K^+\pi^-$	$\pi \leftrightarrow \mu$	$2996 < m_{(\pi\to\mu)\mu} < 3196$ and (π`IsMuon` or $\pi\text{DLL}_{\mu\pi} > 0$)
	$K \leftrightarrow \mu$	$2996 < m_{(K\to\mu)\mu} < 3196$ and (K`IsMuon` or $K\text{DLL}_{\mu\pi} > 0$)
$B^0 \to \psi(2S)K^+\pi^-$	$\pi \leftrightarrow \mu$	$3626 < m_{(\pi\to\mu)\mu} < 3746$ and (π`IsMuon` or $\pi\text{DLL}_{\mu\pi} > 5$)
	$K \leftrightarrow \mu$	$3626 < m_{(K\to\mu)\mu} < 3746$ and (K`IsMuon` or $K\text{DLL}_{\mu\pi} > 5$)
$B^0 \to K^+\pi^-\mu^+\mu^-$	$K \leftrightarrow \pi$	$(K\text{DLL}_{K\pi} - \pi\text{DLL}_{K\pi}) > 10$

band. They should also have a $K^+\mu^+\mu^-$ invariant mass, $m_{K\mu\mu}$, consistent with the nominal B^+ mass. Although these candidates do not accumulate in the signal region, they have a different angular distribution to the combinatorial background and could therefore bias the measurements of the angular moments, described in Sec. 7.8.

Candidates from $B^0 \to J/\psi\, K^+\pi^-$ and $B^0 \to \psi(2S)K^+\pi^-$ decays can contribute a background if the π^- (K^+) is misidentified as a μ^- (μ^+) and the μ^- (μ^+) is misidentified as a π^- (K^+). For the case of $\mu^- \leftrightarrow \pi^-$ swaps, the invariant mass of the π^- and the μ^+, after assigning the π^- the μ mass hypothesis, should be consistent with the known J/ψ or $\psi(2S)$ masses. Likewise, for the case of $K^+ \leftrightarrow \mu^+$ swaps, the invariant mass of the K^+ and the μ^-, after assigning the K^+ the μ mass hypothesis, should be consistent with the known J/ψ or $\psi(2S)$ masses. These new mass hypotheses are denoted $m_{(\pi\to\mu)\mu}$ and $m_{(K\to\mu)\mu}$ respectively.

A background contribution from genuine $B^0 \to K^+\pi^-\mu^+\mu^-$ decays can also be formed when the two hadron hypotheses are swapped. This leads to B^0 candidates being incorrectly reconstructed as $\overline{B}^0$ candidates and vice versa. These candidates are vetoed using hadron identification criteria.

7.4 Agreement between data and simulation

Good agreement between data and simulated decays is necessary in order to be able to accurately model the distortion of the angular distributions caused by the trigger, reconstruction and selection. The acceptance correction, described in detail in Sec. 7.5, is determined from simulated four-body $B^0 \to K^+\pi^-\mu^+\mu^-$ decays generated according to a phase space distribution. Data driven techniques are used to improve the agreement between data and simulation. The PID distributions in simulation are corrected using a method known as 'resampling'. To take into account remaining differences, simulated decays are reweighted to match specific distributions in data.

7.4.1 PID resampling

Particle indentification information is used in two places within the selection of $B^0 \to K^+\pi^-\mu^+\mu^-$: to veto peaking backgrounds and as input to the multivariate classifier. The PID distributions are known to disagree between data and simulation. In order to improve the agreement, the distribution of each PID variable is resampled in a two stage process. Firstly, histograms of the PID variables are produced in bins of the number of tracks in the event, pseudorapidity and p_{T} using calibration samples in data. These samples include $D^{*+} \to D^0(\to K^-\pi^+)\pi^-$, $\Lambda \to p\pi^-$ and $J/\psi \to \mu^+\mu^-$ decays. Secondly, the PID variables for simulated decay are updated using the corresponding histogram as a probability density function to draw a new value. This sampled value replaces the PID variable for the simulated candidate and is used in subsequent operations. The validation of the method for the variables $K\mathrm{DLL}_{K\pi}$ and $\pi\mathrm{DLL}_{K\pi}$ is shown in Fig. 38 using background-subtracted[6] $B^0 \to J/\psi K^{*0}$ decays in data and simulated $B^0 \to J/\psi K^{*0}$ decays. The distributions for the remaining PID variables used in the candidate selection are shown in Appendix B.1.

7.4.2 Reweighting candidates to account for residual differences

The distribution of three variables that show differences between data and simulation are used to derive a candidate reweighting to improve agreement. These three variables are the following: the number of tracks in the event, the p_{T} of the B^0 candidate and the B^0 vertex quality χ^2/ndof.

[6] The background component is subtracted using the *sPlot* technique [79].

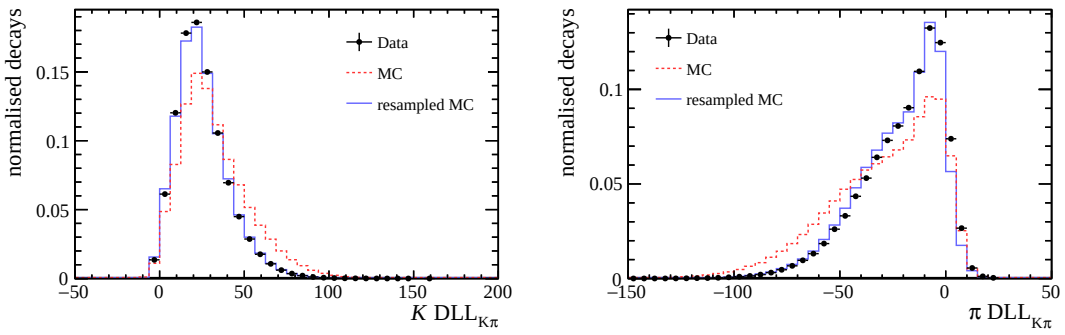


Figure 38: Data-simulation agreement for the PID variables $K\text{DLL}_{K\pi}$ and $\pi\text{DLL}_{K\pi}$. The black data points show the distributions for background-subtracted $B^0 \to J/\psi\, K^{*0}$ decays in data. The red dashed histograms show the nominal distribution for simulated $B^0 \to J/\psi\, K^{*0}$ decays. The blue histograms show the distribution for simulated $B^0 \to J/\psi\, K^{*0}$ decays after the resampling procedure.

The candidate weights are derived by comparing background-subtracted $B^0 \to J/\psi\, K^{*0}$ decays in data and simulated $B^0 \to J/\psi\, K^{*0}$ decays. The weights are determined sequentially, with the previous weight being applied before the subsequent weight is derived. The candidate weights are then applied to all simulated samples. The validation of the method is shown in Fig. 39. Figure 40 shows the agreement for the BDT response before and after applying the candidate reweighting.

7.5 Acceptance correction

The triggering, reconstruction and selection of signal candidates distorts their kinematic distributions. The dominant acceptance effects are due to the requirements on track momentum and impact parameter.

In order to take into account acceptance effects, a five dimensional efficiency function is determined from simulated four-body $B^0 \to K^+\pi^-\mu^+\mu^-$ decays generated according to a phase space distribution. If the distributions of q^2, $\cos\theta_\ell$, $\cos\theta_K$, ϕ and $m(K^+\pi^-)$ were all generated flat, the distributions of the variables after reconstruction and selection would give the shape of the efficiency. While this is true for $\cos\theta_\ell$, $\cos\theta_K$ and ϕ, it is not the case for q^2 and $m(K^+\pi^-)$. Therefore, the simulated candidates are reweighted in order to transform the reconstructed distributions to the efficiency shape.

The efficiency is parameterised in terms of Legendre polynomials of order n,

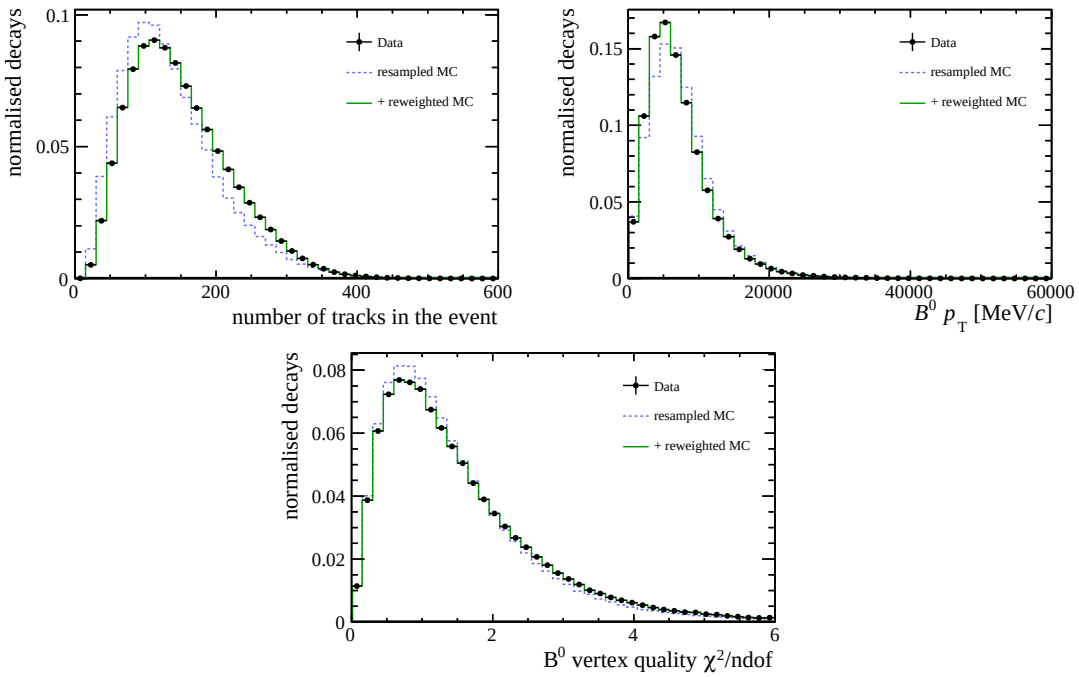


Figure 39: Data-simulation agreement for the variables used to determine the candidate weights. The black data points show the distributions for background-subtracted $B^0 \to J/\psi K^{*0}$ decays in data. The blue dashed histograms show the distribution for resampled, simulated $B^0 \to J/\psi K^{*0}$ decays. The green histograms show the distribution for resampled, simulated $B^0 \to J/\psi K^{*0}$ decays with the candidate weights applied.

$L_n(x)$, as

$$\varepsilon(q^{2\prime}, \cos\theta_\ell, \cos\theta_K, \phi', m'(K^+\pi^-)) = \sum_{hijkl} c_{hijkl}\ L_h(q^{2\prime}) L_i(\cos\theta_\ell) L_j(\cos\theta_K) L_k(\phi') L_l(m'(K^+\pi^-)). \tag{7.3}$$

As the polynomials are defined over the domain $x \in [-1, 1]$, the variables $q^{2\prime}$, ϕ' and $m'(K^+\pi^-)$ are used, which are obtained by linearly transforming q^2, ϕ and $m(K^+\pi^-)$ to lie in this range.

The coefficients c_{hijkl} are determined using a moment analysis of simulated four-body $B^0 \to K^+\pi^-\mu^+\mu^-$ phase-space decays as

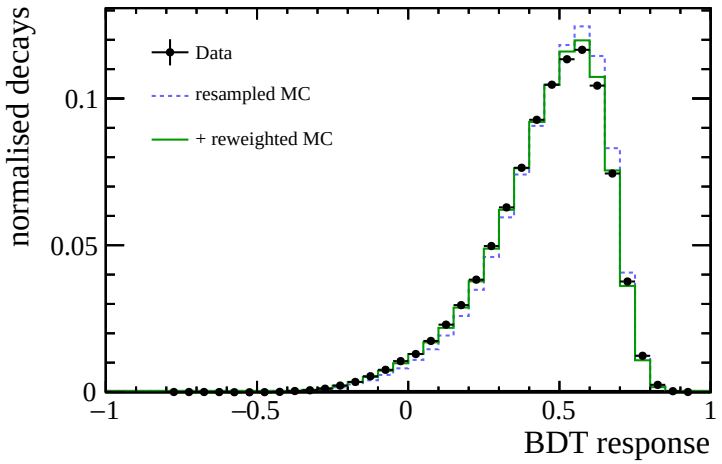


Figure 40: Data-simulation agreement of the BDT response. The black data points show the distributions for background-subtracted $B^0 \to J/\psi\, K^{*0}$ decays in data. The blue dashed histograms show the distribution for resampled, simulated $B^0 \to J/\psi\, K^{*0}$ decays. The green histograms show the distribution for resampled, simulated $B^0 \to J/\psi\, K^{*0}$ decays with the candidate weights applied.

$$c_{hijkl} = \frac{1}{\sum w_n} \sum_{n=0}^{N} w_n \left(\frac{2h+1}{2}\right)\left(\frac{2i+1}{2}\right)\left(\frac{2j+1}{2}\right)\left(\frac{2k+1}{2}\right)\left(\frac{2l+1}{2}\right) \times L_h(q^{2\prime}) L_i(\cos\theta_\ell) L_j(\cos\theta_K) L_k(\phi') L_l(m'(K^+\pi^-)) \,, \tag{7.4}$$

where w_n is the per-candidate weight taking into account both the non-flat distributions of q^2 and $m(K^+\pi^-)$, and the candidate weights described in Sec. 7.4.2. The factors of $(2a+1)/2$ arise from the orthogonality of the Legendre polynomials,

$$\int_{-1}^{+1} L_a(x) L_{a'}(x) \mathrm{d}x = \frac{2}{2a+1} \delta_{aa'} \,. \tag{7.5}$$

The sum in Eq. 7.3 encompasses $L_n(x)$ up to fourth order in $\cos\theta_\ell$ and $m'(K^+\pi^-)$, sixth order in ϕ' and $q^{2\prime}$, and eighth order in $\cos\theta_K$. The order of polynomial used in each case is the lowest order possible that gives good agreement between the efficiency function and the simulated four-body $B^0 \to K^+\pi^-\mu^+\mu^-$ phase-space decays. The angular acceptance as a function of each of the kinematic variables in the region $0.1 < q^2 < 10.0\,\mathrm{GeV}^2/c^4$, $795 < m(K^+\pi^-) < 1530\,\mathrm{MeV}/c^2$ is shown in Fig. 41.

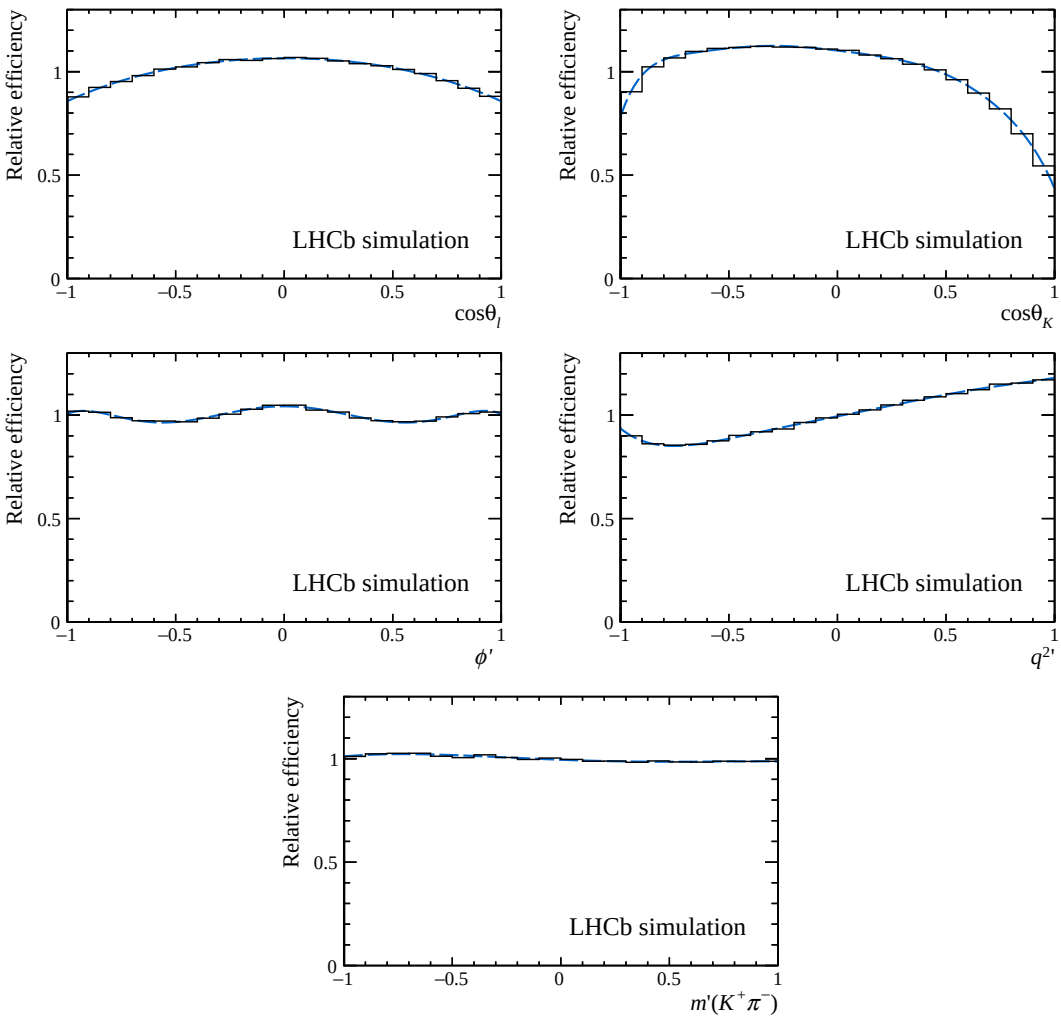


Figure 41: Relative efficiency in each of the kinematic variables in the region $0.1 < q^2 < 10.0\,\text{GeV}^2/c^4$, $795 < m(K^+\pi^-) < 1530\,\text{MeV}/c^2$ as determined from a moment analysis of simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ decays, shown as a histogram. The efficiency function is shown by the blue dashed line.

7.6 The $m(K^+\pi^-\mu^+\mu^-)$ invariant mass distribution

The $m(K^+\pi^-\mu^+\mu^-)$ invariant mass is used to discriminate between signal and combinatorial background. The signal distribution is modelled using the sum of two Gaussian functions with a common mean, each with a power-law tail on the low-mass side. The combinatorial background is modelled using an exponential function. The parameters describing the shape of the mass distribution of the signal are determined from a fit to the $B^0 \to J/\psi\, K^*(892)^0$ decay in data, as shown in Fig. 42, and are subsequently fixed when fitting the $B^0 \to K^+\pi^-\mu^+\mu^-$

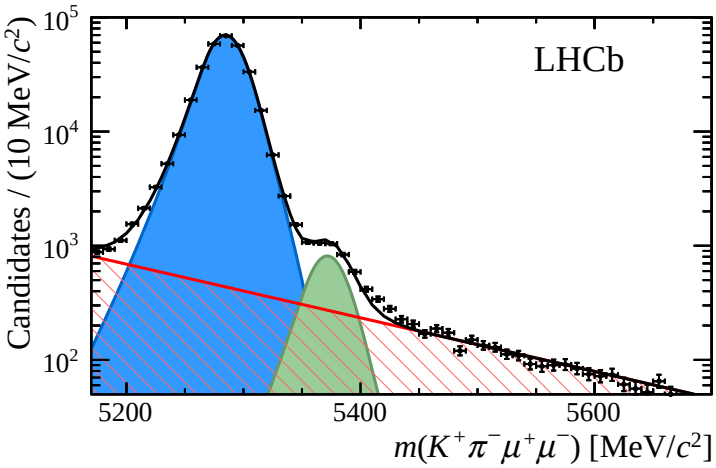


Figure 42: Invariant mass $m(K^+\pi^-\mu^+\mu^-)$ for the control decay $B^0 \to J/\psi\, K^*(892)^0$. The solid black line represents the total fitted function. The individual components of the $B^0 \to J/\psi\, K^{*0}$ (blue shaded area), $B_s^0 \to J/\psi\, \overline{K}^{*0}$ (green shaded area) and combinatorial background (red hatched area) are also shown.

candidates. An additional component is included to model the contribution from $B_s^0 \to J/\psi\, \overline{K}^{*0}$ in the fit to the control mode.

A single scaling factor is used to correct the width of the Gaussian functions to account for variations in the shape of the mass distribution of the signal observed in simulation, due to the different regions of $m(K^+\pi^-)$ and q^2 between the control mode and signal mode. This factor is determined by first fitting the $m(K^+\pi^-\mu^+\mu^-)$ distribution for simulated four-body $B^0 \to K^+\pi^-\mu^+\mu^-$ decays in the $B^0 \to J/\psi\, K^*(892)^0$ region. All the fit parameters are then fixed, except for s_σ which is allowed to float in the subsequent fits to the $m(K^+\pi^-\mu^+\mu^-)$ distribution in each of the q^2 bins in the $B^0 \to K^+\pi^-\mu^+\mu^-$ signal region. The distribution of s_σ as a function of q^2 in the $1330 < m(K^+\pi^-) < 1530\,\text{MeV}/c^2$ region is shown in Fig. 43.

In order to cross-check the method, the scaling factor is determined both for simulated four-body $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and for data in the region $9.22 < q^2 < 9.96\,\text{GeV}^2/c^4$ and $1330 < m(K^+\pi^-) < 1530\,\text{MeV}/c^2$. The value of s_σ determined from simulation is in good agreement with that determined from data.

The fit to $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in the range $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$ is shown in Fig. 44. The individual fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in each of the q^2 bins used for the differential branching fraction measurement are shown in Appendix D. The signal and background yields in each of the q^2 bins, over the

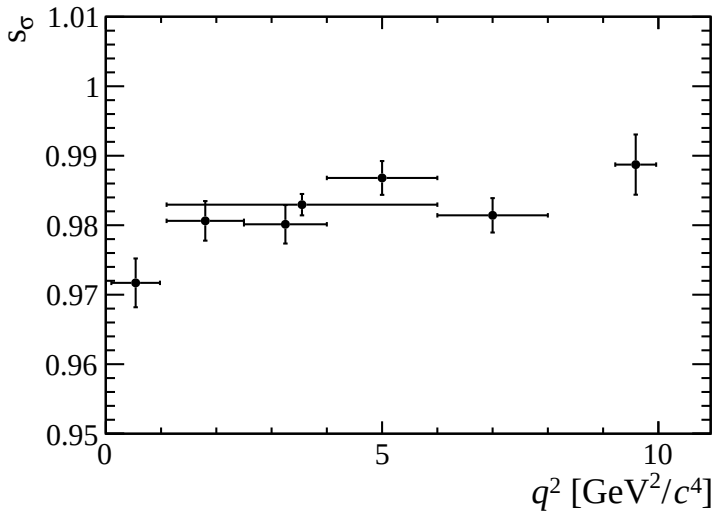


Figure 43: Scaling factor s_σ in bins of q^2 for candidates in the $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 region.

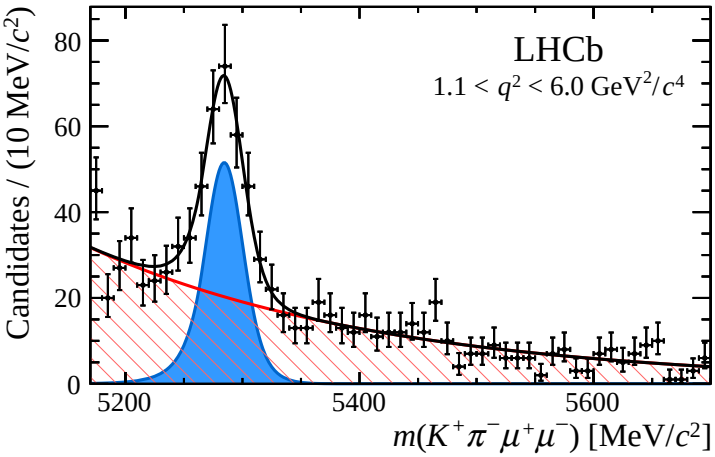


Figure 44: Invariant mass $m(K^+\pi^-\mu^+\mu^-)$ for the signal decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the range $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$. The solid black line represents the total fitted function. The individual components of the signal (blue shaded area) and combinatorial background (red hatched area) are also shown.

range $5170 < m(K^+\pi^-\mu^+\mu^-) < 5700$ MeV/c^2, are given in Table 13.

7.7 Differential branching fraction

The differential branching fraction $\mathrm{d}\mathcal{B}/\mathrm{d}q^2$ of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in an interval $(q^2_{\text{min}}, q^2_{\text{max}})$ is given by

q^2 [GeV/c^2]	Signal yield	Background yield
$[0.10, 0.98]$	67 ± 10	93 ± 11
$[1.10, 2.50]$	80 ± 12	160 ± 15
$[2.50, 4.00]$	75 ± 12	213 ± 17
$[4.00, 6.00]$	75 ± 13	334 ± 21
$[6.00, 8.00]$	60 ± 14	476 ± 25
$[1.10, 6.00]$	229 ± 21	708 ± 31

Table 13: Signal and background yields of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in each of the q^2 bins.

$$\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}q^2} = \frac{1}{(q^2_{\text{max}} - q^2_{\text{min}})} f_{K^*(892)^0} \mathcal{B}(B^0 \to J/\psi\, K^*(892)^0) \mathcal{B}(J/\psi \to \mu^+\mu^-) \times \mathcal{B}(K^*(892)^0 \to K^+\pi^-) \frac{N'_{K^+\pi^-\mu^+\mu^-}}{(1 - F_{\text{S}}^{J/\psi\, K^{*0}}) N'_{J/\psi\, K^{*0}}}, \tag{7.6}$$

where $N'_{K^+\pi^-\mu^+\mu^-}$ and $N'_{J/\psi\, K^{*0}}$ are the acceptance-corrected yields of the $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to J/\psi\, K^{*0}$ decays, respectively. The $B^0 \to J/\psi\, K^{*0}$ yield has to be corrected for the S-wave fraction within the narrow $m(K^+\pi^-)$ window of $B^0 \to J/\psi\, K^{*0}$ decays, $F_{\text{S}}^{J/\psi\, K^{*0}}$. The value of $F_{\text{S}}^{J/\psi\, K^{*0}}$ is obtained from Ref. [85] and is recalculated for the $m(K^+\pi^-)$ range $796 < m(K^+\pi^-) < 996$ MeV/c^2. The branching fractions $\mathcal{B}(B^0 \to J/\psi\, K^*(892)^0)$, $\mathcal{B}(J/\psi \to \mu^+\mu^-)$ and $\mathcal{B}(K^*(892)^0 \to K^+\pi^-)$ are $(1.19 \pm 0.01 \pm 0.08) \times 10^{-3}$ [86], $(5.961 \pm 0.033) \times 10^{-2}$ [4] and 2/3, respectively. The fraction $f_{K^*(892)^0}$ is used to scale the value of $\mathcal{B}(B^0 \to J/\psi\, K^*(892)^0)$ to the appropriate $m(K^+\pi^-)$ range and is calculated by integrating the $K^*(892)^0$ line shape given in Ref. [86] over the range $796 < m(K^+\pi^-) < 996$ MeV/c^2.

7.7.1 Acceptance corrected yields

To avoid making any assumptions about the unknown distributions of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates, the event-by-event efficiencies described in Sec. 7.5 are used to correct the measured yields by calculating the average acceptance weight, where each weight is the reciprocal of the event-by-event efficiency.

For the case where there are only signal candidates present, the average weight would simply be calculated as,

$$\overline{w} = \frac{1}{N}\sum_{i}^{N} w_i, \tag{7.7}$$

where w_i is the event-by-event acceptance weight and N is the number of candidates. An estimate for the error on the average weight is given by,

$$\delta_{\overline{w}} = \sqrt{\frac{1}{N(N-1)}\sum_{i}^{N}(w_i - \overline{w})^2}. \tag{7.8}$$

Due to the presence of background, the average weight calculated in the signal region will be an admixture of the average weight for both signal candidates ($\overline{w}_{sig}$) and background candidates ($\overline{w}_{bkg}$),

$$\overline{w}_{mix} = \frac{N_{sig}\overline{w}_{sig} + N_{bkg}\overline{w}_{bkg}}{N_{sig} + N_{bkg}}, \tag{7.9}$$

where N_{sig} and N_{bkg} are the number of signal and background events in the signal region, respectively. This can be rearranged to give the average weight for the signal candidates,

$$\overline{w}_{sig} = \frac{(N_{sig} + N_{bkg})\overline{w}_{mix} - N_{bkg}\overline{w}_{bkg}}{N_{sig}}. \tag{7.10}$$

However, what is needed for both $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to J/\psi\, K^{*0}$ is the acceptance corrected yield $\overline{w}_{sig}N_{sig}$. This is given by,

$$\overline{w}_{sig}N_{sig} = (N_{sig} + N_{bkg})\overline{w}_{mix} - N_{bkg}\overline{w}_{bkg} \tag{7.11}$$

where the errors are propagated as,

$$\begin{aligned}\sigma^2_{\overline{w}_{sig}N_{sig}} = \; & (N_{sig} + N_{bkg})^2\sigma^2_{\overline{w}_{mix}} + (-N_{bkg})^2\sigma^2_{\overline{w}_{bkg}} \\ & + (\overline{w}_{mix})^2\sigma^2_{N_{sig}} + (\overline{w}_{mix} - \overline{w}_{bkg})^2\sigma^2_{N_{bkg}}.\end{aligned} \tag{7.12}$$

The signal region is defined as $5230 < m(K^+\pi^-\mu^+\mu^-) < 5330$ MeV/c^2 and the background region as $5350 < m(K^+\pi^-\mu^+\mu^-) < 5700$ MeV/c^2. For the resonant mode, the background region is altered to $5450 < m(K^+\pi^-\mu^+\mu^-) < 5700$ MeV/c^2 in order to prevent any potential pollution from $B^0 \to J/\psi\, K^{*0}$ or $B^0_s \to J/\psi\, \overline{K}^{*0}$ candidates.

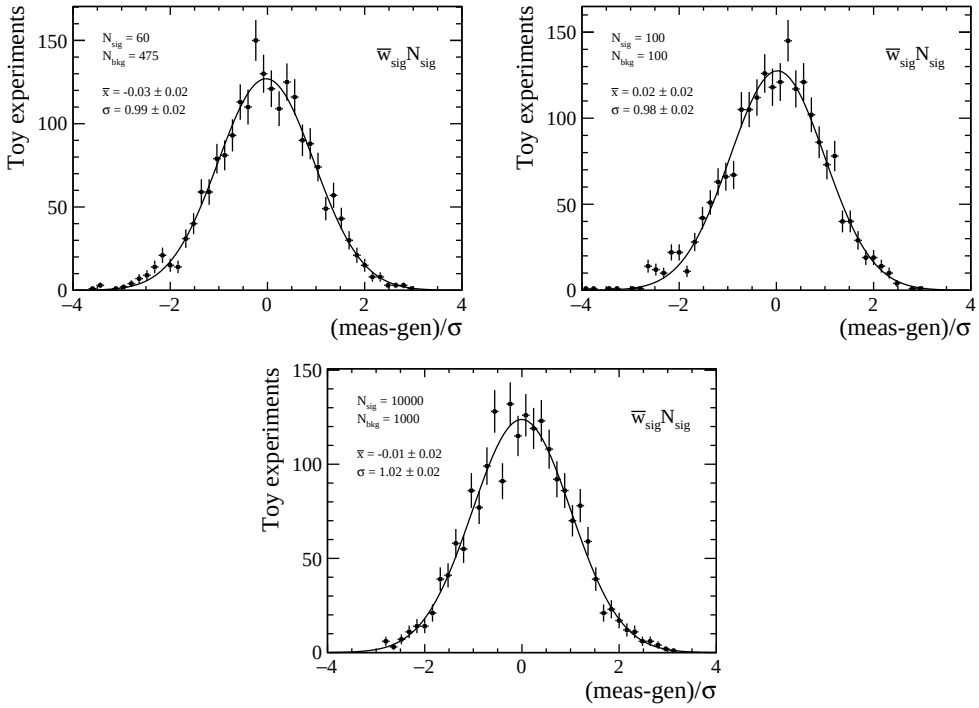


Figure 45: Pull plots for the extraction of $\overline{w}_{sig}N_{sig}$ with different numbers of signal and background candidates.

7.7.2 Toy studies

Toy studies are performed for the extraction of $\overline{w}_{sig}N_{sig}$ with different numbers of signal and background candidates. In each toy, N_{sig} and N_{bkg} are Poisson fluctuated. The nominal mass model, described in Sec. 7.6, is used to generate signal and background candidates. The weights for both signal and background are sampled from two gaussian functions with different means.

Pull studies are performed to check that the method is unbiased and correctly estimates the uncertainty. The pull of the observable a is defined as $(a_{\text{meas}} - a_{\text{gen}})/\sigma(a)_{\text{meas}}$ [87]. If the method is unbiased and the coverage is correct, the pull distribution should be a Gaussian centred at zero with unit width. The pulls for the extraction of $\overline{w}_{sig}N_{sig}$ are shown in Fig. 45. No bias is observed and the statistical error is correctly evaluated.

7.7.3 Results

The results for the differential branching fraction are given in Fig. 46. The uncertainties shown are the quadratic sum of the statistical and systematic uncer-

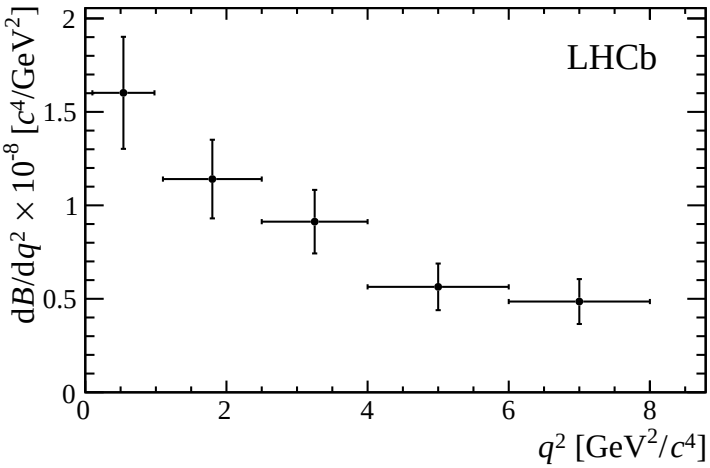


Figure 46: Differential branching fraction of $B^0 \to K^+\pi^-\mu^+\mu^-$ in bins of q^2. The uncertainties shown are the quadratic sum of the statistical and systematic uncertainties.

Table 14: Differential branching fraction of $B^0 \to K^+\pi^-\mu^+\mu^-$ in bins of q^2. The first uncertainty is statistical, the second systematic and the third due to the uncertainty on the $B^0 \to J/\psi K^*(892)^0$ and $J/\psi \to \mu^+\mu^-$ branching fractions.

q^2 [GeV^2/c^4]	$\mathrm{d}\mathcal{B}/\mathrm{d}q^2 \times 10^{-8}$ [c^4/GeV^2]
$[0.10, 0.98]$	$1.60 \pm 0.28 \pm 0.04 \pm 0.11$
$[1.10, 2.50]$	$1.14 \pm 0.19 \pm 0.03 \pm 0.08$
$[2.50, 4.00]$	$0.91 \pm 0.16 \pm 0.03 \pm 0.06$
$[4.00, 6.00]$	$0.56 \pm 0.12 \pm 0.02 \pm 0.04$
$[6.00, 8.00]$	$0.49 \pm 0.11 \pm 0.01 \pm 0.03$
$[1.10, 6.00]$	$0.82 \pm 0.09 \pm 0.02 \pm 0.06$

tainties. The results are also presented in Table 14. The various sources of the systematic uncertainties are described in Sec. 7.9.

7.8 Angular moments analysis

The angular observables defined in Sec. 7.2 are determined using a moments analysis of the angular distribution, as outlined in Ref. [80]. This approach has the advantage of producing stable measurements with well-defined uncertainties even for small data samples.

The 41 background-subtracted and acceptance-corrected moments are esti-

mated as

$$\Gamma_i = \sum_{k=1}^{n_{\text{sig}}} w_k f_i(\Omega_k) - x \sum_{k=1}^{n_{\text{bkg}}} w_k f_i(\Omega_k) \tag{7.13}$$

and the corresponding covariance matrix is estimated as

$$C_{ij} = \sum_{k=1}^{n_{\text{sig}}} w_k^2 f_i(\Omega_k) f_j(\Omega_k) + x^2 \sum_{k=1}^{n_{\text{bkg}}} w_k^2 f_i(\Omega_k) f_j(\Omega_k). \tag{7.14}$$

Here n_{sig} and n_{bkg} correspond to the candidates in the signal and background regions, respectively. The signal region is defined within $\pm 50\,\text{MeV}/c^2$ of the mean B^0 mass, and the background region in the range $5350 < m(K^+\pi^-\mu^+\mu^-) < 5700$ MeV/c^2. The scale factor x is the ratio of the estimated number of background candidates in the signal region over the number of candidates in the background region and is used to normalise the background subtraction. It has been checked in data that the angular distribution of the background is independent of $m(K^+\pi^-\mu^+\mu^-)$ within the precision of this measurement, and that the uncertainty on x has negligible impact on the results. The weights, w_k, are the reciprocals of the candidates' efficiencies and account for the acceptance, described in Sec. 7.5.

The covariance matrix describing the statistical uncertainties on the 40 normalised moments is computed as

$$\overline{C}_{ij} = \left[C_{ij} + \frac{\Gamma_i \Gamma_j}{\Gamma_1^2} C_{11} - \frac{\Gamma_i C_{1j} + \Gamma_j C_{1i}}{\Gamma_1} \right] \frac{1}{\Gamma_1^2}, \quad i, j \in \{2, ..., 41\}. \tag{7.15}$$

7.8.1 Toy studies

In order to validate the moments analysis, toy studies are performed using simulated datasets. Signal decays are generated according to a realistic model taking into account contributions from the $K^*(1410)^0$ [88], $K_0^*(1430)^0$ [89] and $K_2^*(1430)^0$ [90] states. The generated values for each of the moments are calculated using a numeric integration. The specified $m(K^+\pi^-)$ and q^2 range is iterated over and the amplitudes at each point are calculated. The moments are then derived from these amplitudes using the relations given in Appendix A.

The toy datasets are generated in the range $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 and $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$. Each dataset contains a realistic amount of signal and background candidates, the numbers of which are both Poisson fluctuated. Background candidates are generated flat in q^2, $\cos\theta_\ell$, $\cos\theta_K$, ϕ, and $m(K^+\pi^-)$, and

Table 15: Results of the pull studies for toys datasets.

$\overline{\Gamma}_i$	mean	width	$\overline{\Gamma}_i$	mean	width
$\overline{\Gamma}_2$	-0.01 ± 0.02	0.99 ± 0.01	$\overline{\Gamma}_{22}$	0.01 ± 0.02	0.96 ± 0.02
$\overline{\Gamma}_3$	0.01 ± 0.02	0.96 ± 0.02	$\overline{\Gamma}_{23}$	-0.01 ± 0.02	0.96 ± 0.02
$\overline{\Gamma}_4$	-0.03 ± 0.02	0.95 ± 0.02	$\overline{\Gamma}_{24}$	-0.05 ± 0.02	0.98 ± 0.02
$\overline{\Gamma}_5$	-0.03 ± 0.02	1.00 ± 0.02	$\overline{\Gamma}_{25}$	-0.01 ± 0.02	0.99 ± 0.02
$\overline{\Gamma}_6$	0.01 ± 0.02	1.01 ± 0.02	$\overline{\Gamma}_{26}$	-0.05 ± 0.02	0.96 ± 0.02
$\overline{\Gamma}_7$	0.02 ± 0.02	1.01 ± 0.02	$\overline{\Gamma}_{27}$	-0.01 ± 0.02	0.96 ± 0.02
$\overline{\Gamma}_8$	-0.01 ± 0.02	0.97 ± 0.02	$\overline{\Gamma}_{28}$	-0.01 ± 0.02	0.99 ± 0.02
$\overline{\Gamma}_9$	-0.01 ± 0.02	0.99 ± 0.02	$\overline{\Gamma}_{29}$	0.03 ± 0.02	0.97 ± 0.02
$\overline{\Gamma}_{10}$	-0.03 ± 0.02	0.97 ± 0.02	$\overline{\Gamma}_{30}$	0.02 ± 0.02	1.02 ± 0.02
$\overline{\Gamma}_{11}$	0.01 ± 0.02	0.94 ± 0.02	$\overline{\Gamma}_{31}$	-0.00 ± 0.02	0.97 ± 0.01
$\overline{\Gamma}_{12}$	-0.03 ± 0.02	0.98 ± 0.02	$\overline{\Gamma}_{32}$	0.00 ± 0.02	1.01 ± 0.02
$\overline{\Gamma}_{13}$	0.00 ± 0.02	0.98 ± 0.02	$\overline{\Gamma}_{33}$	-0.03 ± 0.02	0.98 ± 0.02
$\overline{\Gamma}_{14}$	0.03 ± 0.02	0.95 ± 0.01	$\overline{\Gamma}_{34}$	-0.00 ± 0.02	0.98 ± 0.02
$\overline{\Gamma}_{15}$	-0.01 ± 0.02	0.97 ± 0.01	$\overline{\Gamma}_{35}$	-0.01 ± 0.02	0.98 ± 0.02
$\overline{\Gamma}_{16}$	0.00 ± 0.02	0.97 ± 0.02	$\overline{\Gamma}_{36}$	-0.00 ± 0.02	0.96 ± 0.01
$\overline{\Gamma}_{17}$	0.03 ± 0.02	1.01 ± 0.02	$\overline{\Gamma}_{37}$	-0.01 ± 0.02	0.97 ± 0.02
$\overline{\Gamma}_{18}$	-0.00 ± 0.02	0.97 ± 0.02	$\overline{\Gamma}_{38}$	0.02 ± 0.02	0.95 ± 0.02
$\overline{\Gamma}_{19}$	-0.01 ± 0.02	1.00 ± 0.02	$\overline{\Gamma}_{39}$	0.01 ± 0.02	0.96 ± 0.01
$\overline{\Gamma}_{20}$	-0.00 ± 0.02	0.97 ± 0.02	$\overline{\Gamma}_{40}$	0.02 ± 0.02	0.99 ± 0.02
$\overline{\Gamma}_{21}$	0.01 ± 0.02	0.97 ± 0.02	$\overline{\Gamma}_{41}$	-0.01 ± 0.02	0.96 ± 0.02

according to an exponential distribution in $m(K^+\pi^-\mu^+\mu^-)$. Acceptance effects are also included using the efficiency function described in Sec. 7.5.

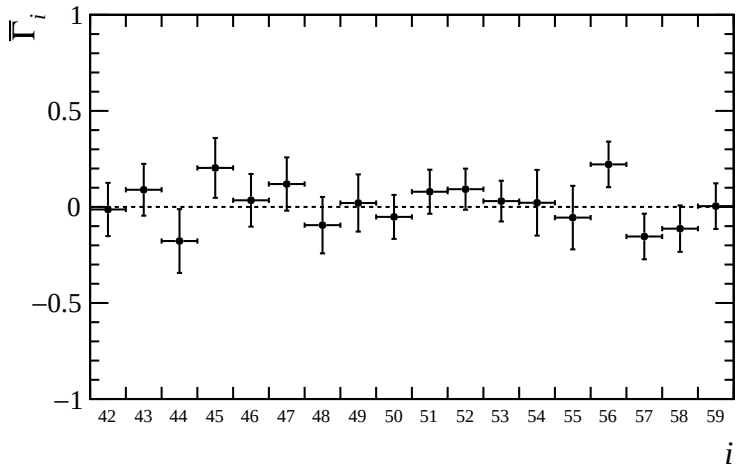
The 40 normalised moments $\overline{\Gamma}_2 - \overline{\Gamma}_{41}$ are measured for each of the toy datasets. Pull studies are performed to check that the method is unbiased and correctly estimates the uncertainty. The results of the pull studies are shown in Table 15 and in Appendix E. No bias is observed for any of the measured moments and the statistical errors are correctly evaluated.

7.8.2 F-wave moments

The angular analysis assumes that the $K^+\pi^-$ system is in a S-, P- or D-wave configuration only. Contributions from F-wave states (or higher) are assumed to be negligible. This assumption is cross-checked by measuring the 18 moments corresponding to the $K^+\pi^-$ in a F-wave configuration. The measured values are shown in Fig. 47 and Table 16. All the measured moments are consistent with

zero.

Figure 47: Results of the cross-check of the F-wave contribution. All the measured moments are consistent with zero.



	$f_i(\Omega)$	Value
$\overline{\Gamma}_{42}$	$P_5^0 Y_0^0$	-0.01 ± 0.14
$\overline{\Gamma}_{43}$	$P_6^0 Y_0^0$	0.09 ± 0.13
$\overline{\Gamma}_{44}$	$P_5^0 Y_2^0$	-0.18 ± 0.17
$\overline{\Gamma}_{45}$	$P_6^0 Y_2^0$	0.20 ± 0.16
$\overline{\Gamma}_{46}$	$P_5^1 \sqrt{2}\, Re(Y_2^1)$	0.03 ± 0.14
$\overline{\Gamma}_{47}$	$P_6^1 \sqrt{2}\, Re(Y_2^1)$	0.12 ± 0.14
$\overline{\Gamma}_{48}$	$P_5^1 \sqrt{2}\, Im(Y_2^1)$	-0.09 ± 0.15
$\overline{\Gamma}_{49}$	$P_6^1 \sqrt{2}\, Im(Y_2^1)$	0.02 ± 0.15
$\overline{\Gamma}_{50}$	$P_5^0 \sqrt{2}\, Re(Y_2^2)$	-0.05 ± 0.11
$\overline{\Gamma}_{51}$	$P_6^0 \sqrt{2}\, Re(Y_2^2)$	0.08 ± 0.11
$\overline{\Gamma}_{52}$	$P_5^0 \sqrt{2}\, Im(Y_2^2)$	0.09 ± 0.11
$\overline{\Gamma}_{53}$	$P_6^0 \sqrt{2}\, Im(Y_2^2)$	0.03 ± 0.11
$\overline{\Gamma}_{54}$	$P_5^0 Y_1^0$	0.02 ± 0.17
$\overline{\Gamma}_{55}$	$P_6^0 Y_1^0$	-0.06 ± 0.17
$\overline{\Gamma}_{56}$	$P_5^1 \sqrt{2}\, Re(Y_1^1)$	0.22 ± 0.12
$\overline{\Gamma}_{57}$	$P_6^1 \sqrt{2}\, Re(Y_1^1)$	-0.15 ± 0.12
$\overline{\Gamma}_{58}$	$P_5^1 \sqrt{2}\, Im(Y_1^1)$	-0.11 ± 0.12
$\overline{\Gamma}_{59}$	$P_6^1 \sqrt{2}\, Im(Y_1^1)$	0.00 ± 0.12

Table 16: Results of the cross-check of the F-wave contribution. All the measured moments are consistent with zero.

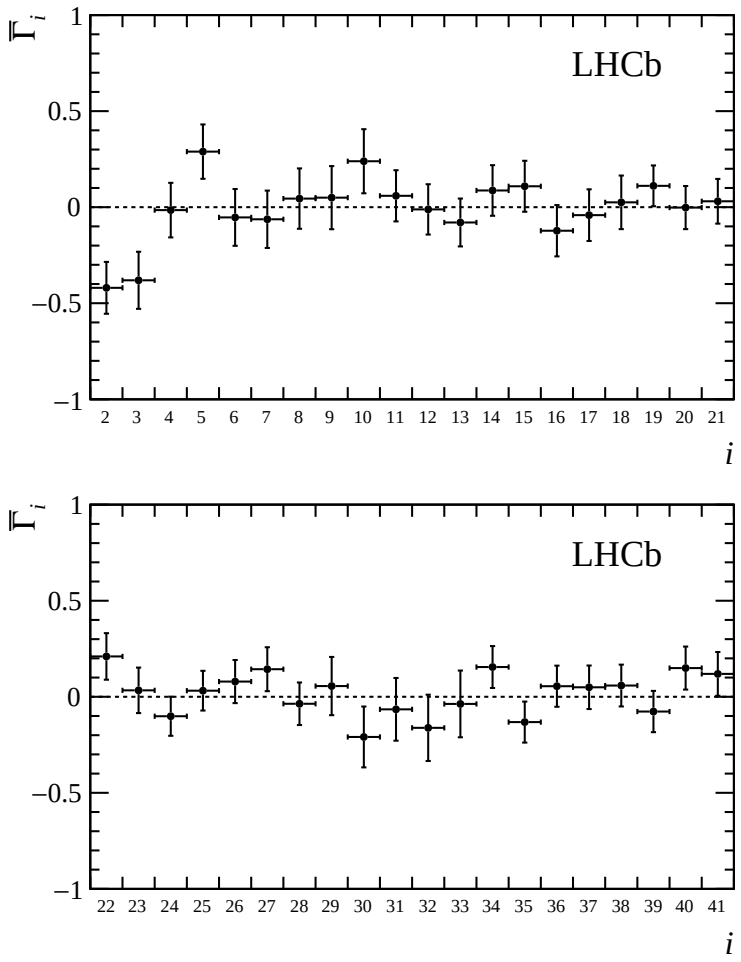


Figure 48: Measurement of the normalised moments, $\overline{\Gamma}_i$, of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the range $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$ and $1330 < m(K^+\pi^-) < 1530\,\text{MeV}/c^2$. The uncertainties shown are the quadratic sum of the statistical and systematic uncertainties.

7.8.3 Results

The results for the normalised moments, $\overline{\Gamma}_i$, are given in Fig. 48. The uncertainties shown are the sums in quadrature of the statistical and systematic uncertainties. The results are also presented in Table 17. The various sources of the systematic uncertainties are described in Sec. 7.9.

The distributions of each of the decay angles within the signal region are shown in Fig. 49. The estimated signal distribution is derived from the moments model by evaluating the sum in Eq. 7.1, which is found to provide a good representation of the data for each of the decay angles.

Table 17: Measurement of the normalised moments, $\overline{\Gamma}_i$, of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the range $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ and $1330 < m(K^+\pi^-) < 1530\,\mathrm{MeV}/c^2$. The first uncertainty is statistical and the second systematic. The 0.00 values for the systematic uncertainty of certain moments is simply due to the number of significant digits shown.

$\overline{\Gamma}_i$	Value	$\overline{\Gamma}_i$	Value
$\overline{\Gamma}_2$	$-0.42 \pm 0.13 \pm 0.03$	$\overline{\Gamma}_{22}$	$0.21 \pm 0.12 \pm 0.01$
$\overline{\Gamma}_3$	$-0.38 \pm 0.15 \pm 0.01$	$\overline{\Gamma}_{23}$	$0.03 \pm 0.12 \pm 0.01$
$\overline{\Gamma}_4$	$-0.02 \pm 0.14 \pm 0.01$	$\overline{\Gamma}_{24}$	$-0.10 \pm 0.10 \pm 0.01$
$\overline{\Gamma}_5$	$0.29 \pm 0.14 \pm 0.02$	$\overline{\Gamma}_{25}$	$0.03 \pm 0.10 \pm 0.01$
$\overline{\Gamma}_6$	$-0.05 \pm 0.14 \pm 0.04$	$\overline{\Gamma}_{26}$	$0.08 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_7$	$-0.06 \pm 0.15 \pm 0.03$	$\overline{\Gamma}_{27}$	$0.14 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_8$	$0.04 \pm 0.16 \pm 0.01$	$\overline{\Gamma}_{28}$	$-0.04 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_9$	$0.05 \pm 0.16 \pm 0.02$	$\overline{\Gamma}_{29}$	$0.06 \pm 0.15 \pm 0.04$
$\overline{\Gamma}_{10}$	$0.24 \pm 0.17 \pm 0.02$	$\overline{\Gamma}_{30}$	$-0.21 \pm 0.15 \pm 0.04$
$\overline{\Gamma}_{11}$	$0.06 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{31}$	$-0.07 \pm 0.16 \pm 0.01$
$\overline{\Gamma}_{12}$	$-0.01 \pm 0.13 \pm 0.02$	$\overline{\Gamma}_{32}$	$-0.16 \pm 0.17 \pm 0.02$
$\overline{\Gamma}_{13}$	$-0.08 \pm 0.12 \pm 0.01$	$\overline{\Gamma}_{33}$	$-0.04 \pm 0.17 \pm 0.02$
$\overline{\Gamma}_{14}$	$0.09 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{34}$	$0.15 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{15}$	$0.11 \pm 0.13 \pm 0.00$	$\overline{\Gamma}_{35}$	$-0.13 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{16}$	$-0.12 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{36}$	$0.05 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{17}$	$-0.04 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{37}$	$0.05 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{18}$	$0.03 \pm 0.14 \pm 0.01$	$\overline{\Gamma}_{38}$	$0.06 \pm 0.11 \pm 0.00$
$\overline{\Gamma}_{19}$	$0.11 \pm 0.11 \pm 0.01$	$\overline{\Gamma}_{39}$	$-0.08 \pm 0.11 \pm 0.00$
$\overline{\Gamma}_{20}$	$-0.00 \pm 0.11 \pm 0.01$	$\overline{\Gamma}_{40}$	$0.15 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{21}$	$0.03 \pm 0.12 \pm 0.01$	$\overline{\Gamma}_{41}$	$0.12 \pm 0.11 \pm 0.01$

The D-wave fraction, F_D, is estimated using the moments $\overline{\Gamma}_5$ and $\overline{\Gamma}_{10}$ as

$$F_\mathrm{D} \equiv -\frac{7}{18}\left(2\overline{\Gamma}_5 + 5\sqrt{5}\overline{\Gamma}_{10}\right). \tag{7.16}$$

Naively, one would expect a large D-wave contribution in this region, as was seen in the amplitude analysis of $B^0 \to J/\psi\, K^+\pi^-$ [86]. However, in $B^0 \to K^+\pi^-\mu^+\mu^-$ no significant D-wave contribution is seen and, with the limited statistics currently available, it is only possible to set an upper limit of $F_\mathrm{D} < 0.29$ at 95% confidence level. This might be an indication of a large breaking of QCD factorisation in this decay mode. Additionally, the values of the moments $\overline{\Gamma}_2$ and $\overline{\Gamma}_3$ imply the presence of large interference effects between the S- and P- or D-wave contributions.

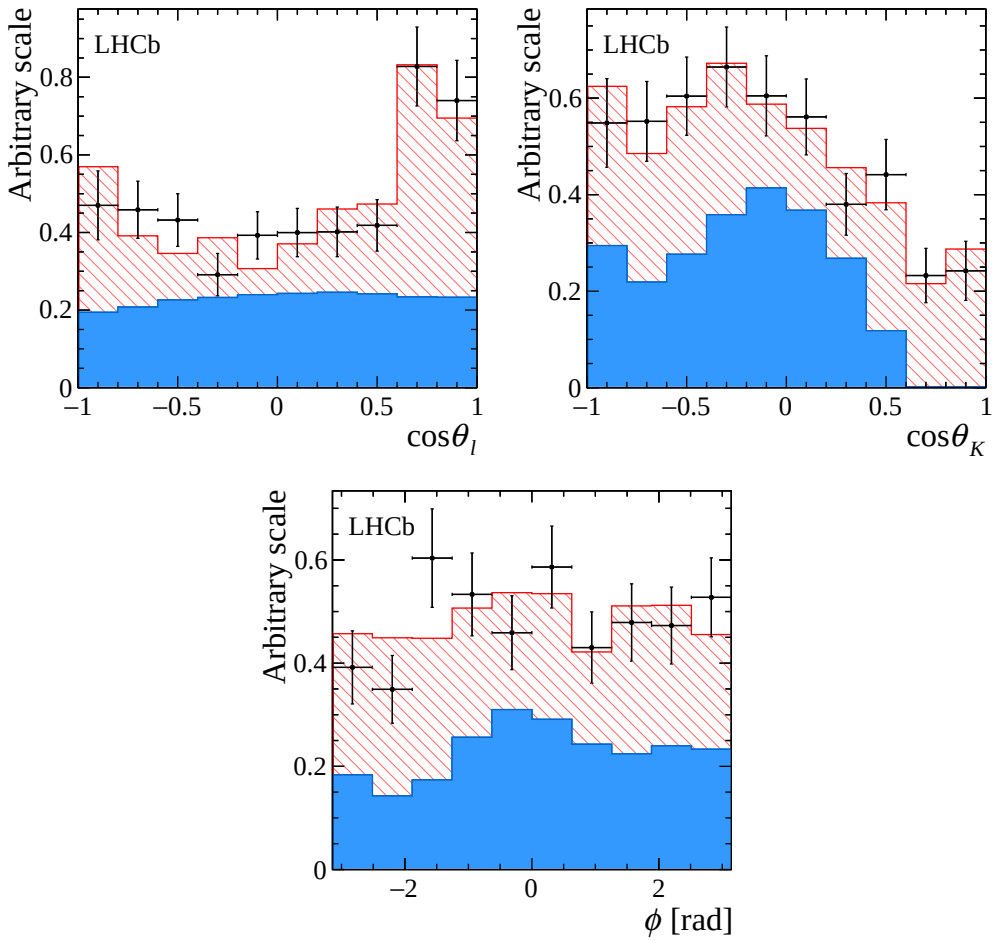


Figure 49: The distributions of each of the decay angles within the signal region. The acceptance-corrected data is represented by the points with error bars. The estimated signal distribution is shown by the blue shaded histogram. The projected background from the upper mass sideband is shown by the red hatched histogram, which is stacked onto the signal histogram.

7.9 Systematic uncertainties

The main sources of systematic uncertainty for the differential branching fraction and the angular moments analysis are described in detail below and summarised in Table 18. The systematic uncertainties are significantly smaller than the statistical uncertainties.

The differential branching fraction and angular moments analysis share several common systematic effects: the statistical uncertainty on the acceptance function due to the size of the simulated sample from which it is determined, differences between data candidates and the simulated candidates used for the determination

Table 18: Summary of the main sources of systematic uncertainty for the differential branching fraction and the angular moments analysis. Typical ranges are quoted for the different q^2 bins used in the differential branching fraction measurement and the different moments measured in the angular moments analysis. The systematic uncertainties are significantly smaller than the statistical uncertainties.

Source	$d\mathcal{B}/dq^2 \times 10^{-8}$ $[c^4/\,\mathrm{GeV}^2]$	$\overline{\Gamma}_i$
Acceptance stat. uncertainty	0.006−0.030	0.003−0.013
Data-simulation differences	0.001−0.014	0.001−0.007
Peaking backgrounds	0.013−0.026	0.001−0.040
$\mathcal{B}(B^0 \to J/\psi\, K^*(892)^0)$	0.033−0.110	−

of the acceptance function and contributions from residual peaking background candidates. The differential branching fraction also has a systematic uncertainty due to the uncertainty on the branching fraction of the decay $B^0 \to J/\psi\, K^*(892)^0$. This is the dominant systematic uncertainty for the differential branching fraction and is shown separately in Table 14.

The size of the systematic uncertainties associated with the determination of the acceptance correction and residual peaking background contributions are evaluated using pseudo-experiments, where samples are generated varying one or more parameters. The differential branching fraction and each of the moments are evaluated using both the nominal model and the systematically varied model. In general, the systematic uncertainty is then taken as the average of the difference between the two models over a large number of pseudo-experiments. The exception to this is the statistical uncertainty of the acceptance function. In order to account for this statistical variation, the standard deviation of the difference between the two models is used instead.

The effect of the statistical uncertainty on the acceptance function is evaluated using pseudoexperiments where candidates are generated with an acceptance that is varied according to the covariance matrix for the moments of the acceptance function. The effect of differences between the data candidates and the simulated candidates is evaluated using pseudoexperiments where candidates are generated with an acceptance determined from simulated candidates without applying the data-simulation corrections described in Sec. 7.4.

The effect of residual peaking background contributions is evaluated using pseudoexperiments where peaking background components are generated in addi-

tion to the signal and the combinatorial background. The angular distributions of the peaking backgrounds are taken from data by isolating the decays using specific selections. In Appendix F, the method is shown explicitly for the case of $B^0 \to J/\psi K^+\pi^-$ decays with a $\pi \leftrightarrow \mu$ swap.

All other sources of systematic uncertainties investigated, such as the choice of the $m(K^+\pi^-\mu^+\mu^-)$ signal model and the resolution in the angular variables, are found to have a negligible impact.

7.10 Summary

This thesis presents measurements of the differential branching fraction and angular moments of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^+\pi^-$ invariant mass range $1330 < m(K^+\pi^-) < 1530$ MeV/c^2. The data sample corresponds to an integrated luminosity of $3\,\text{fb}^{-1}$ of pp collision data collected by the LHCb experiment. The differential branching fraction is reported in five narrow q^2 bins between 0.1 and $8.0\,\text{GeV}^2/c^4$ and in the range $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$, where an angular moments analysis is also performed.

The measured values of the angular observables $\overline{\Gamma}_2$ and $\overline{\Gamma}_3$ point towards the presence of large interference effects between the S- and P- or D-wave contributions. Using only $\overline{\Gamma}_5$ and $\overline{\Gamma}_{10}$ it is possible to estimate the D-wave fraction, F_D, yielding an upper limit of $F_\text{D} < 0.29$ at 95% confidence level. This value is lower than naively expected from amplitude analyses of $B^0 \to J/\psi K^+\pi^-$ decays [86].

The underlying Wilson coefficients may be extracted from the normalised moments and covariance matrix presented in this analysis, when combined with a prediction for the form factors. While first estimates for the form factors are given in Ref. [78], no interpretation of the results in terms of the Wilson coefficients is made at this time. With additional input from theory, these results could provide further contributions to understanding the pattern of deviations with respect to SM predictions that has been observed in other $b \to s\mu^+\mu^-$ transitions.

Appendices

A Angular distribution

The transversity-basis moments of the 41 orthonormal angular functions defined in Eq. 7.1 are shown in Table 19. The orthonormal angular basis is constructed out of spherical harmonics, $Y_l^m \equiv Y_l^m(\theta_\ell, \phi)$, and reduced spherical harmonics, $P_l^m \equiv \sqrt{2\pi} Y_l^m(\theta_K, 0)$. The S-, P- and D-wave transversity amplitudes are denoted as $S^{\{L,R\}}$, $H^{\{L,R\}}_{\{0,\parallel,\perp\}}$ and $D^{\{L,R\}}_{\{0,\parallel,\perp\}}$, respectively.

Table 19: The transversity-basis moments of the 41 orthonormal angular functions $f_i(\Omega)$ in Eq. 7.1.

i	$f_i(\Omega)$	$\Gamma_i^{L,\mathrm{tr}}(q^2)$	$\eta_i^{L\to R}$
1	$P_0^0 Y_0^0$	$\left[\lvert H_0^L\rvert^2 + \lvert H_\parallel^L\rvert^2 + \lvert H_\perp^L\rvert^2 + \lvert S^L\rvert^2 + \lvert D_0^L\rvert^2 + \lvert D_\parallel^L\rvert^2 + \lvert D_\perp^L\rvert^2\right]$	$+ (L \to R)$
2	$P_1^0 Y_0^0$	$2\left[\frac{2}{\sqrt{5}} Re(H_0^L D_0^{L*}) + Re(S^L H_0^{L*}) + \sqrt{\frac{3}{5}}\, Re(H_\parallel^L D_\parallel^{L*} + H_\perp^L D_\perp^{L*})\right]$	"
3	$P_2^0 Y_0^0$	$\frac{\sqrt{5}}{7}\,(\lvert D_\parallel^L\rvert^2 + \lvert D_\perp^L\rvert^2)$ - $\frac{1}{\sqrt{5}}\,(\lvert H_\parallel^L\rvert^2 + \lvert H_\perp^L\rvert^2) + \frac{2}{\sqrt{5}}\,\lvert H_0^L\rvert^2 + \frac{10}{7\sqrt{5}}\,\lvert D_0^L\rvert^2 + 2\, Re(S^L D_0^{L*})$	"
4	$P_3^0 Y_0^0$	$\frac{6}{\sqrt{35}}\left[-Re(H_\parallel^L D_\parallel^{L*} + H_\perp^L D_\perp^{L*}) + \sqrt{3} Re(H_0^L D_0^{L*})\right]$	"
5	$P_4^0 Y_0^0$	$\frac{2}{7}\left[-2(\lvert D_\parallel^L\rvert^2 + \lvert D_\perp^L\rvert^2) + 3\lvert D_0^L\rvert^2\right]$	"
6	$P_0^0 Y_2^0$	$\frac{1}{2\sqrt{5}}\left[(\lvert D_\parallel^L\rvert^2 + \lvert D_\perp^L\rvert^2) + (\lvert H_\parallel^L\rvert^2 + \lvert H_\perp^L\rvert^2) - 2\lvert S^L\rvert^2 - 2\lvert D_0^L\rvert^2 - 2\lvert H_0^L\rvert^2\right]$	"
7	$P_1^0 Y_2^0$	$\left[\frac{\sqrt{3}}{5} Re(H_\parallel^L D_\parallel^{L*} + H_\perp^L D_\perp^{L*}) - \frac{2}{\sqrt{5}} Re(S^L H_0^{L*}) - \frac{4}{5} Re(H_0^L D_0^{L*})\right]$	"
8	$P_2^0 Y_2^0$	$\left[\frac{1}{14}(\lvert D_\parallel^L\rvert^2 + \lvert D_\perp^L\rvert^2) - \frac{2}{7}\lvert D_0^L\rvert^2 - \frac{1}{10}(\lvert H_\parallel^L\rvert^2 + \lvert H_\perp^L\rvert^2) - \frac{2}{5}\lvert H_0^L\rvert^2 - \frac{2}{\sqrt{5}} Re(S^L D_0^{L*})\right]$	"
9	$P_3^0 Y_2^0$	$-\frac{3}{5\sqrt{7}}\left[Re(H_\parallel^L D_\parallel^{L*} + H_\perp^L D_\perp^{L*}) + 2\sqrt{3}\, Re(H_0^L D_0^{L*})\right]$	"
10	$P_4^0 Y_2^0$	$-\frac{2}{7\sqrt{5}}\left[\lvert D_\parallel^L\rvert^2 + \lvert D_\perp^L\rvert^2 + 3\lvert D_0^L\rvert^2\right]$	"
11	$P_1^1\sqrt{2}\, Re(Y_2^1)$	$-\frac{3}{\sqrt{10}}\left[\sqrt{\frac{2}{3}} Re(H_\parallel^L S^{L*}) - \sqrt{\frac{2}{15}} Re(H_\parallel^L D_0^{L*}) + \sqrt{\frac{2}{5}} Re(D_\parallel^L H_0^{L*})\right]$	"
12	$P_2^1\sqrt{2}\, Re(Y_2^1)$	$-\frac{3}{5}\left[Re(H_\parallel^L H_0^{L*}) + \sqrt{\frac{5}{3}} Re(D_\parallel^L S^{L*}) + \frac{5}{7\sqrt{3}} Re(D_\parallel^L D_0^{L*})\right]$	"
13	$P_3^1\sqrt{2}\, Re(Y_2^1)$	$-\frac{6}{5\sqrt{14}}\left[2\, Re(D_\parallel^L H_0^{L*}) + \sqrt{3}\, Re(H_\parallel^L D_0^{L*})\right]$	"
14	$P_4^1\sqrt{2}\, Re(Y_2^1)$	$-\frac{6}{7\sqrt{2}} Re(D_\parallel^L D_0^{L*})$	"
15	$P_1^1\sqrt{2}\, Im(Y_2^1)$	$3\left[\frac{1}{\sqrt{15}} Im(H_\perp^L S^{L*}) + \frac{1}{5} Im(D_\perp^L H_0^{L*}) - \frac{1}{5\sqrt{3}} Im(H_\perp^L D_0^{L*})\right]$	"
16	$P_2^1\sqrt{2}\, Im(Y_2^1)$	$3\left[\frac{1}{7\sqrt{3}} Im(D_\perp^L D_0^{L*}) + \frac{1}{5} Im(H_\perp^L H_0^{L*}) + \frac{1}{\sqrt{15}} Im(D_\perp^L S^{L*})\right]$	"
17	$P_3^1\sqrt{2}\, Im(Y_2^1)$	$\frac{6}{5\sqrt{14}}\left[2\, Im(D_\perp^L H_0^{L*}) + \sqrt{3}\, Im(H_\perp^L D_0^{L*})\right]$	"
18	$P_4^1\sqrt{2}\, Im(Y_2^1)$	$\frac{6}{7\sqrt{2}} Im(D_\perp^L D_0^{L*})$	"
19	$P_0^0\sqrt{2}\, Re(Y_2^2)$	$-\frac{3}{2\sqrt{15}}\left[(\lvert H_\parallel^L\rvert^2 - \lvert H_\perp^L\rvert^2) + (\lvert D_\parallel^L\rvert^2 - \lvert D_\perp^L\rvert^2)\right]$	"
20	$P_1^0\sqrt{2}\, Re(Y_2^2)$	$-\frac{3}{5}\left[Re(H_\parallel^L D_\parallel^{L*}) - Re(D_\perp^L H_\perp^{L*})\right]$	"
21	$P_2^0\sqrt{2}\, Re(Y_2^2)$	$\frac{\sqrt{3}}{2}\left[-\frac{1}{7}(\lvert D_\parallel^L\rvert^2 - \lvert D_\perp^L\rvert^2) + \frac{1}{5}(\lvert H_\parallel^L\rvert^2 - \lvert H_\perp^L\rvert^2)\right]$	"
22	$P_3^0\sqrt{2}\, Re(Y_2^2)$	$\frac{3}{5}\sqrt{\frac{3}{7}}\left[Re(H_\parallel^L D_\parallel^{L*}) - Re(D_\perp^L H_\perp^{L*})\right]$	"
23	$P_4^0\sqrt{2}\, Re(Y_2^2)$	$\frac{2}{7}\sqrt{\frac{3}{5}}(\lvert D_\parallel^L\rvert^2 - \lvert D_\perp^L\rvert^2)$	"
24	$P_0^0\sqrt{2}\, Im(Y_2^2)$	$\sqrt{\frac{3}{5}}\left[Im(H_\perp^L H_\parallel^{L*}) + Im(D_\perp^L D_\parallel^{L*})\right]$	"
25	$P_1^0\sqrt{2}\, Im(Y_2^2)$	$\frac{3}{5} Im(H_\perp^L D_\parallel^{L*} + D_\perp^L H_\parallel^{L*})$	"
26	$P_2^0\sqrt{2}\, Im(Y_2^2)$	$\sqrt{3}\left[\frac{1}{7} Im(D_\perp^L D_\parallel^{L*}) - \frac{1}{5} Im(H_\perp^L H_\parallel^{L*})\right]$	"
27	$P_3^0\sqrt{2}\, Im(Y_2^2)$	$-\frac{3}{5}\sqrt{\frac{3}{7}} Im(D_\perp^L H_\parallel^{L*} + H_\perp^L D_\parallel^{L*})$	"
28	$P_4^0\sqrt{2}\, Im(Y_2^2)$	$-\frac{4}{7}\sqrt{\frac{3}{5}} Im(D_\perp^L D_\parallel^{L*})$	"
29	$P_0^0 Y_1^0$	$-\sqrt{3}\left[Re(H_\perp^L H_\parallel^{L*}) + Re(D_\perp^L D_\parallel^{L*})\right]$	- $(L \to R)$
30	$P_1^0 Y_1^0$	$-\frac{3}{\sqrt{5}} Re(H_\perp^L D_\parallel^{L*} + H_\parallel^L D_\perp^{L*})$	"
31	$P_2^0 Y_1^0$	$-\frac{3}{\sqrt{15}}\left[\frac{5}{7} Re(D_\perp^L D_\parallel^{L*}) - Re(H_\perp^L H_\parallel^{L*})\right]$	"
32	$P_3^0 Y_1^0$	$\frac{9}{\sqrt{105}} Re(H_\perp^L D_\parallel^{L*} + H_\parallel^L D_\perp^{L*})$	"
33	$P_4^0 Y_1^0$	$\frac{4\sqrt{3}}{7} Re(D_\perp^L D_\parallel^{L*})$	"
34	$P_1^1\sqrt{2}\, Re(Y_1^1)$	$\sqrt{\frac{3}{5}}\left[\sqrt{5}\, Re(H_\perp^L S^{L*}) + \sqrt{3}\, Re(D_\perp^L H_0^{L*}) - Re(H_\perp^L D_0^{L*})\right]$	"
35	$P_2^1\sqrt{2}\, Re(Y_1^1)$	$3\left[\frac{1}{\sqrt{5}} Re(H_\perp^L H_0^{L*}) + \frac{1}{\sqrt{3}} Re(D_\perp^L S^{L*}) + \frac{5}{21}\sqrt{\frac{3}{5}} Re(D_\perp^L D_0^{L*})\right]$	"
36	$P_3^1\sqrt{2}\, Re(Y_1^1)$	$\frac{6}{\sqrt{70}}\left[2\, Re(D_\perp^L H_0^{L*}) + \sqrt{3}\, Re(H_\perp^L D_0^{L*})\right]$	"
37	$P_4^1\sqrt{2}\, Re(Y_1^1)$	$\frac{3\sqrt{10}}{7} Re(D_\perp^L D_0^{L*})$	"
38	$P_1^1\sqrt{2}\, Im(Y_1^1)$	$-\sqrt{\frac{3}{5}}\left[\sqrt{5}\, Im(H_\parallel^L S^{L*}) + \sqrt{3}\, Im(D_\parallel^L H_0^{L*}) - Im(H_\parallel^L D_0^{L*})\right]$	"
39	$P_2^1\sqrt{2}\, Im(Y_1^1)$	$-\sqrt{\frac{3}{5}}\left[\sqrt{3}\, Im(H_\parallel^L H_0^{L*}) + \sqrt{5}\, Im(D_\parallel^L S^{L*}) + \frac{5}{7} Im(D_\parallel^L D_0^{L*})\right]$	"
40	$P_3^1\sqrt{2}\, Im(Y_1^1)$	$-6\sqrt{\frac{1}{70}}\left[2\, Im(D_\parallel^L H_0^{L*}) + \sqrt{3}\, Im(H_\parallel^L D_0^{L*})\right]$	"
41	$P_4^1\sqrt{2}\, Im(Y_1^1)$	$-\frac{3}{7}\sqrt{10}\, Im(D_\parallel^L D_0^{L*})$	"

B Agreement between data and simulation

B.1 PID resampling

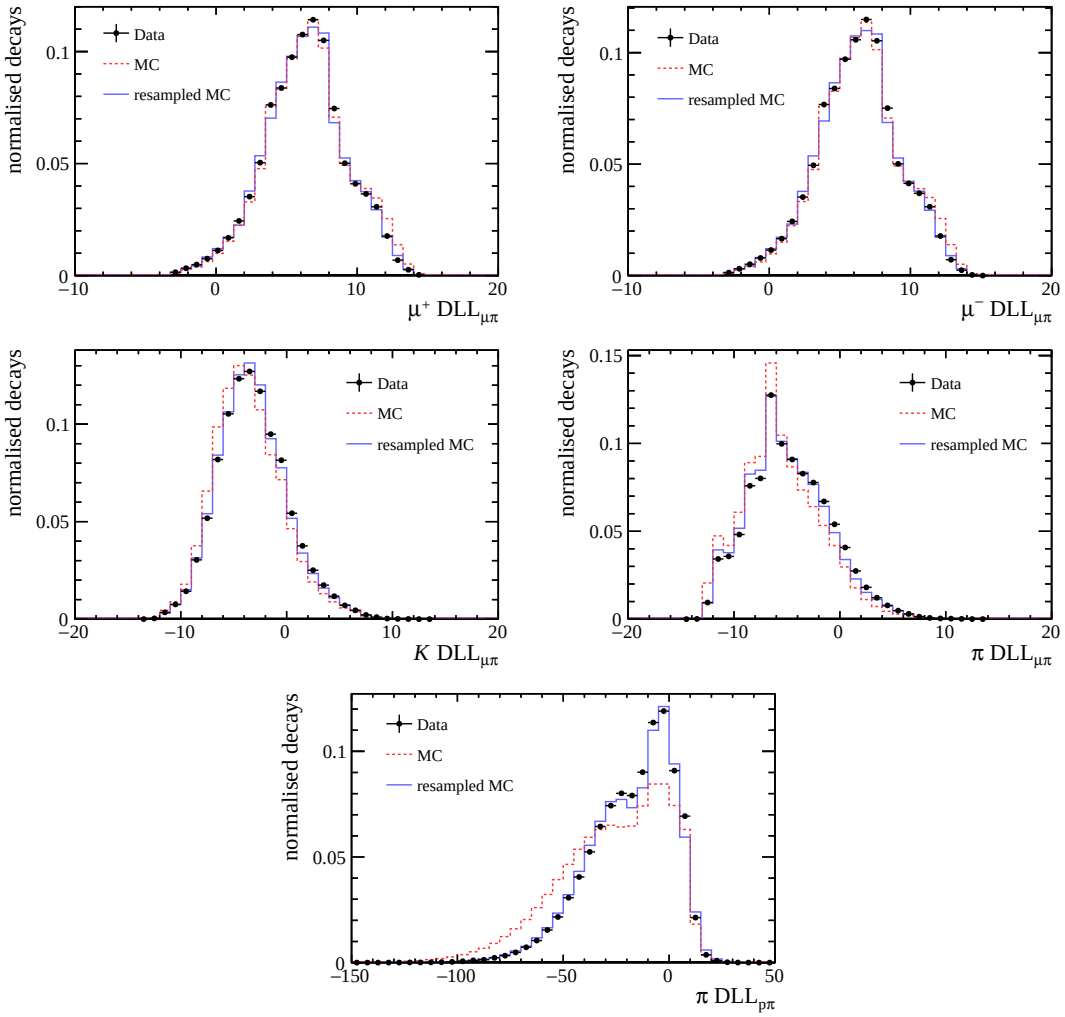


Figure 50: Data-simulation agreement for the PID variables used in the selection of $B^0 \to K^+\pi^-\mu^+\mu^-$. The black data points show the distributions for background-subtracted $B^0 \to J/\psi K^{*0}$ decays in data. The red dashed histograms show the nominal distribution for simulated $B^0 \to J/\psi K^{*0}$ decays. The blue histograms show the distribution for simulated $B^0 \to J/\psi K^{*0}$ decays after the resampling procedure.

B.2 Data-MC agreement for BDT input variables

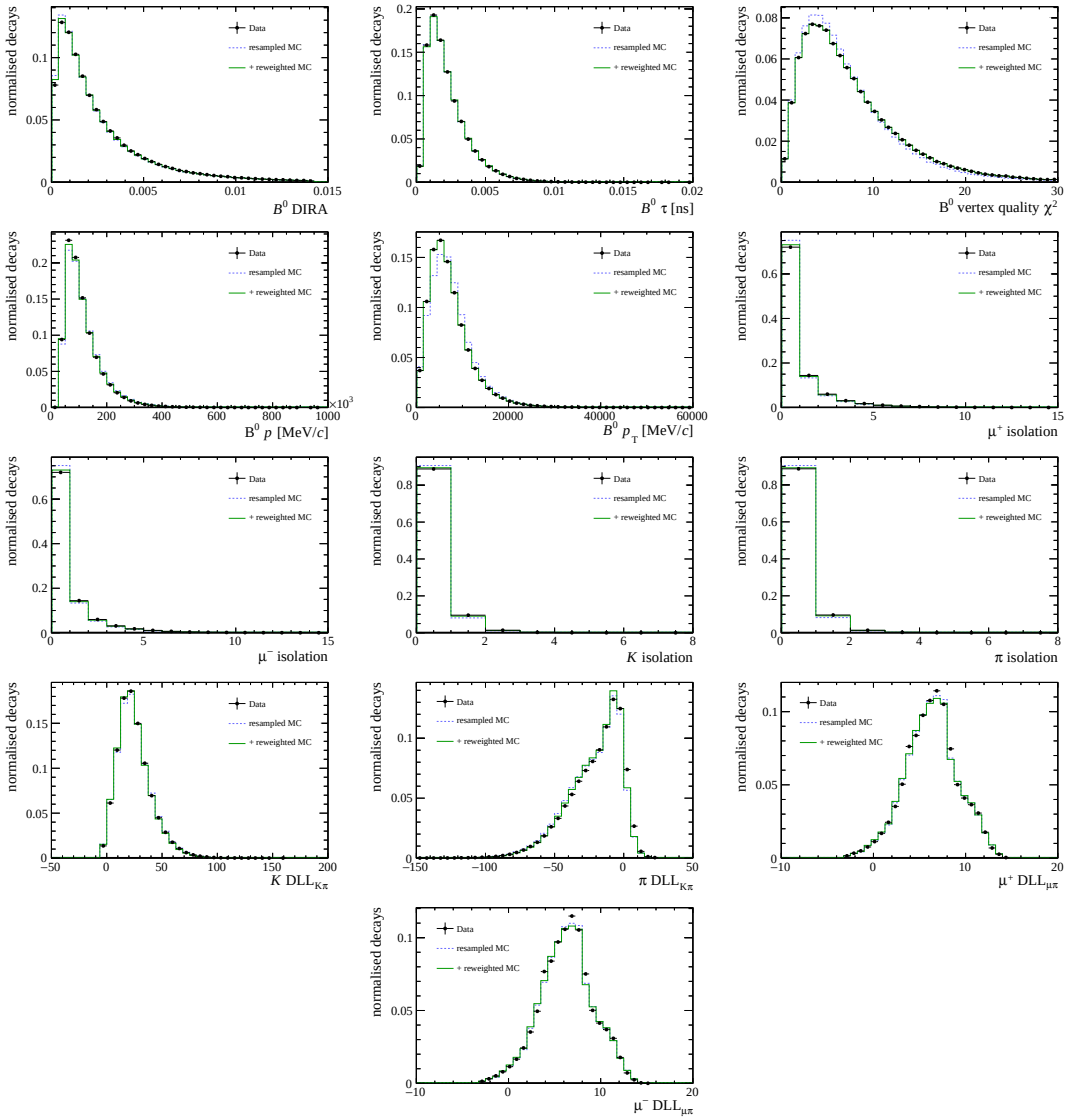


Figure 51: Data-simulation agreement for the variables for each of the variables used as input to the BDT. The black data points show the distributions for background-subtracted $B^0 \to J/\psi K^{*0}$ decays in data. The blue dashed histograms show the distribution for resampled, simulated $B^0 \to J/\psi K^{*0}$ decays. The green histograms show the distribution for resampled, simulated $B^0 \to J/\psi K^{*0}$ decays with the candidate weights applied.

C Acceptance correction

In order to study the correlations between the kinematic variables used in the acceptance correction it is useful to look at two dimensional distributions. Figures 52–55 show the two dimensional distributions for each of the pairs of variables. The plots on the left show the distributions for the simulated decays used to determine the acceptance correction. The plots on the right show the distributions for toy events generated flat in each of the variables and weighted by the efficiency determined from the acceptance parameterisation. Good agreement is found between the two distributions.

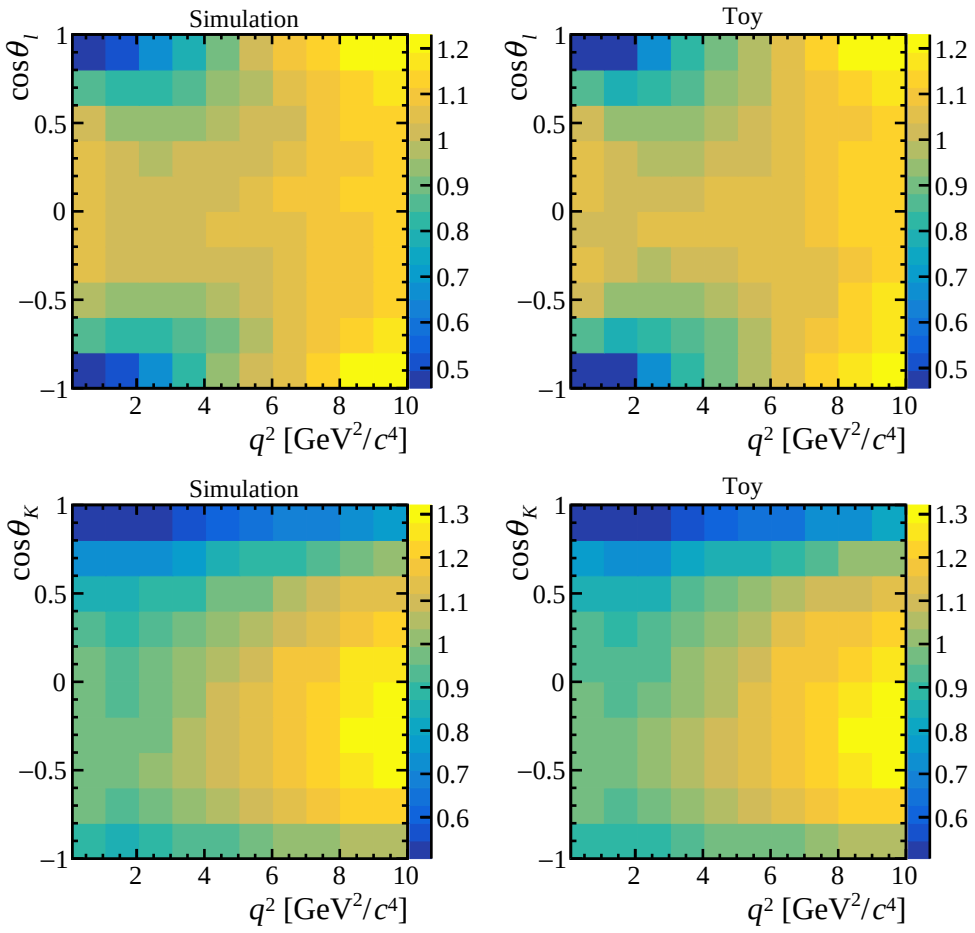


Figure 52: Two dimensional distributions of the acceptance parameterisation.

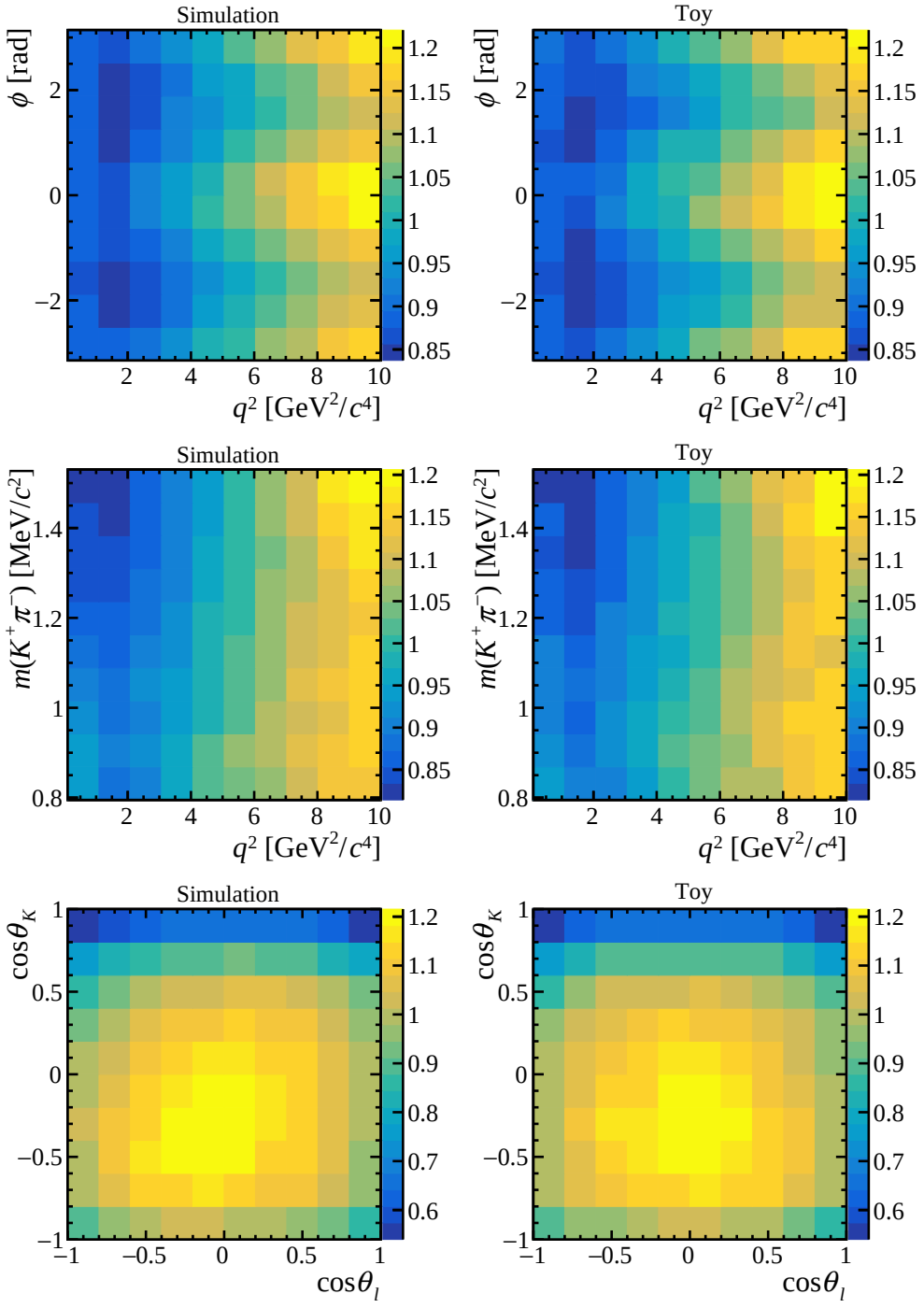


Figure 53: Two dimensional distributions of the acceptance parameterisation.

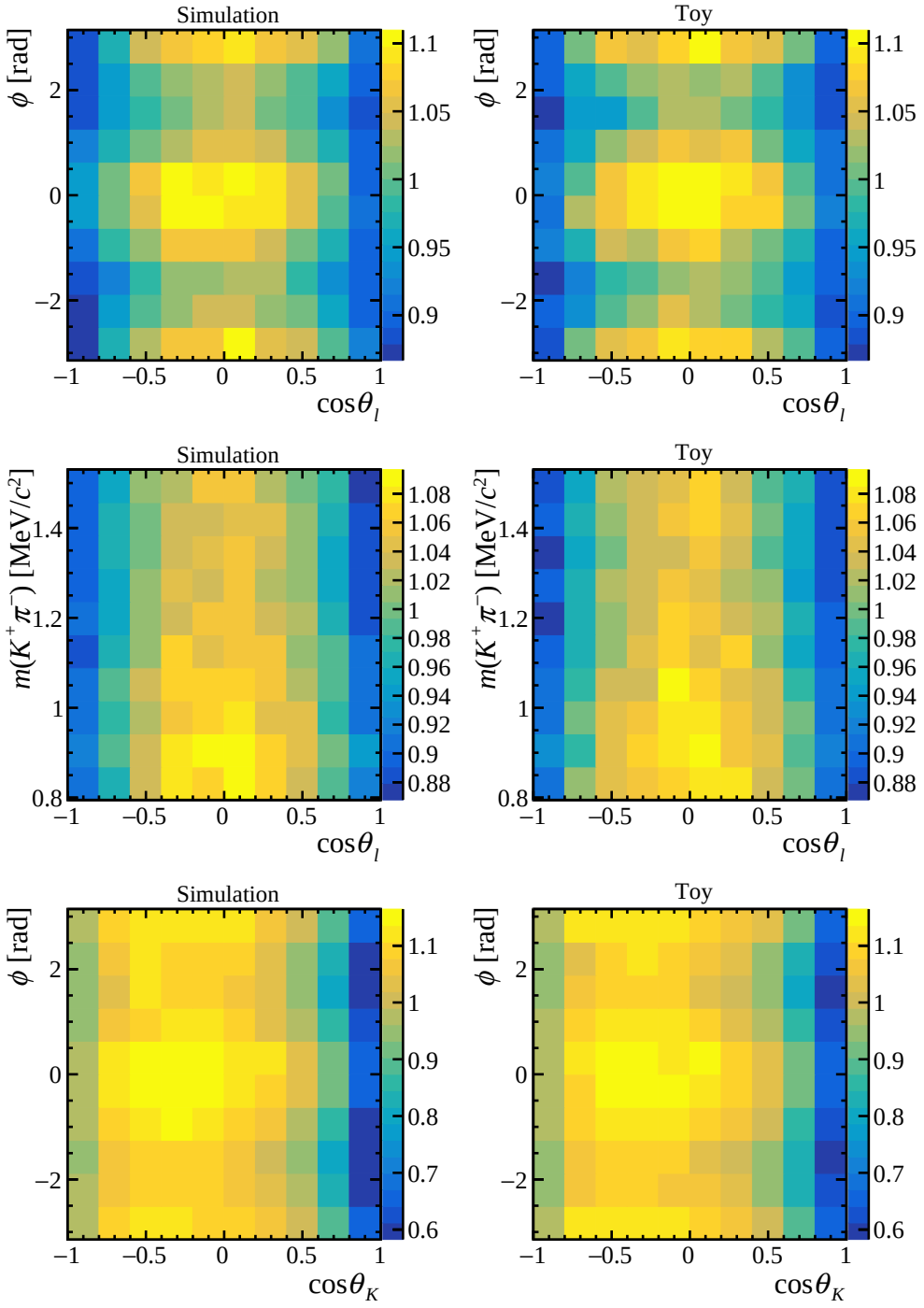


Figure 54: Two dimensional distributions of the acceptance parameterisation.

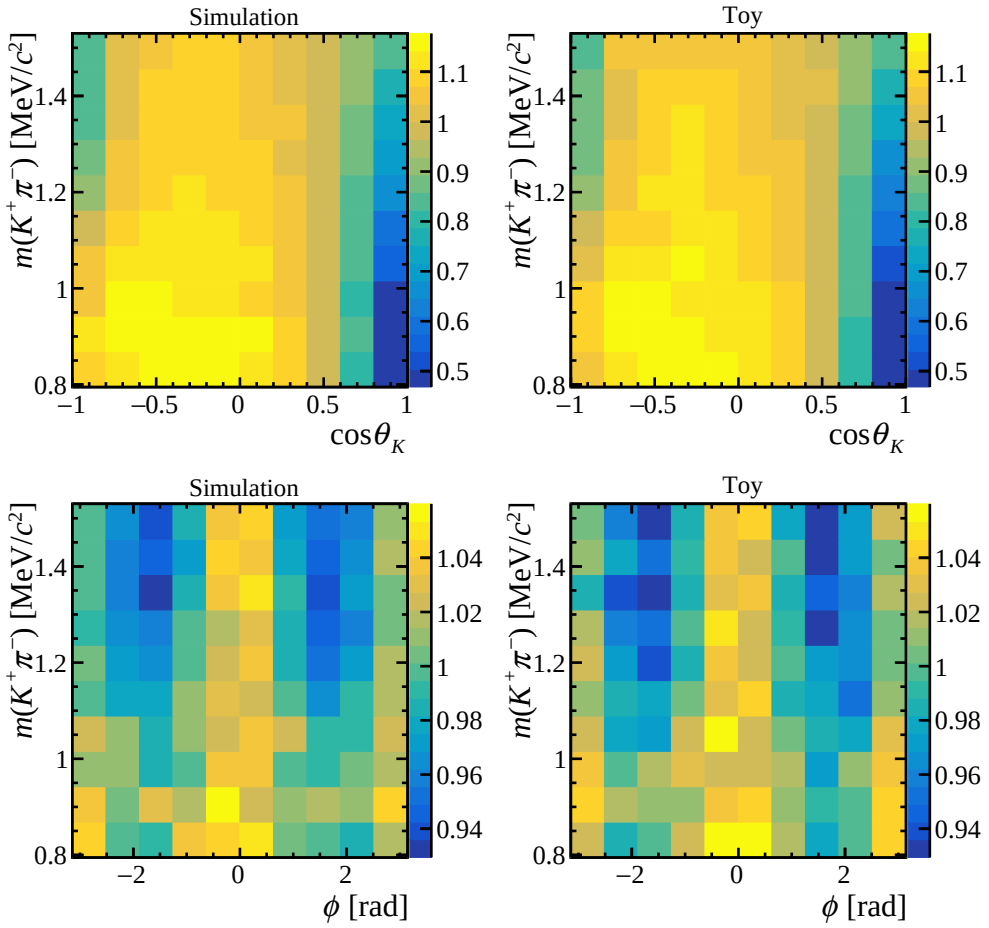


Figure 55: Two dimensional distributions of the acceptance parameterisation.

D The $m(K^+\pi^-\mu^+\mu^-)$ invariant mass distribution

Figure 56 shows the fits to the $m(K^+\pi^-\mu^+\mu^-)$ distribution in each of the q^2 bins used for the differential branching fraction measurement.

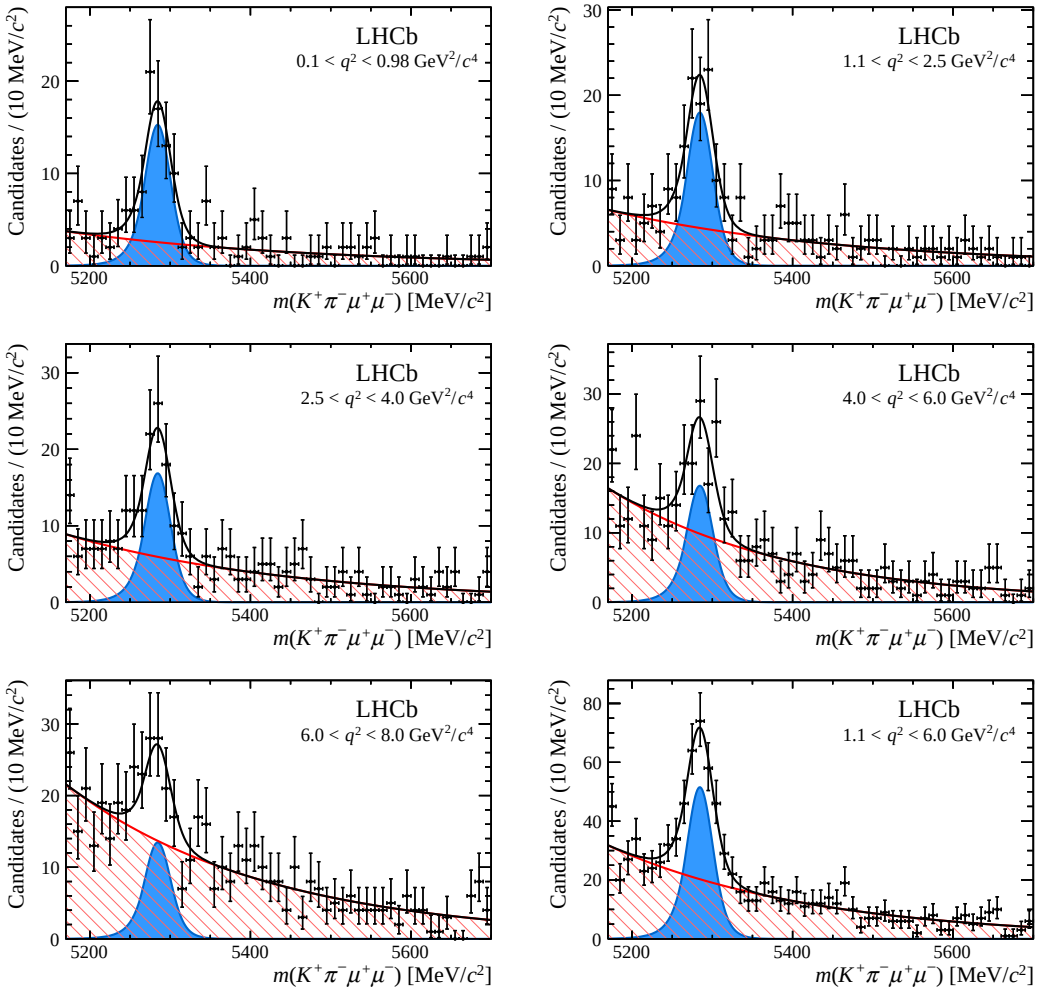


Figure 56: Invariant mass $m(K^+\pi^-\mu^+\mu^-)$ for the signal decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in each of the q^2 bins used for the differential branching fraction measurement. The solid black line represents the total fitted function. The individual components of the signal (blue shaded area) and combinatorial background (red hatched area) are also shown.

E Toy studies for the angular moments analysis

The results of the pull studies described in Sec. 7.8.1 are shown in Fig. 57 and Fig. 58. No bias is observed for any of the measured moments and the statistical errors are correctly evaluated.

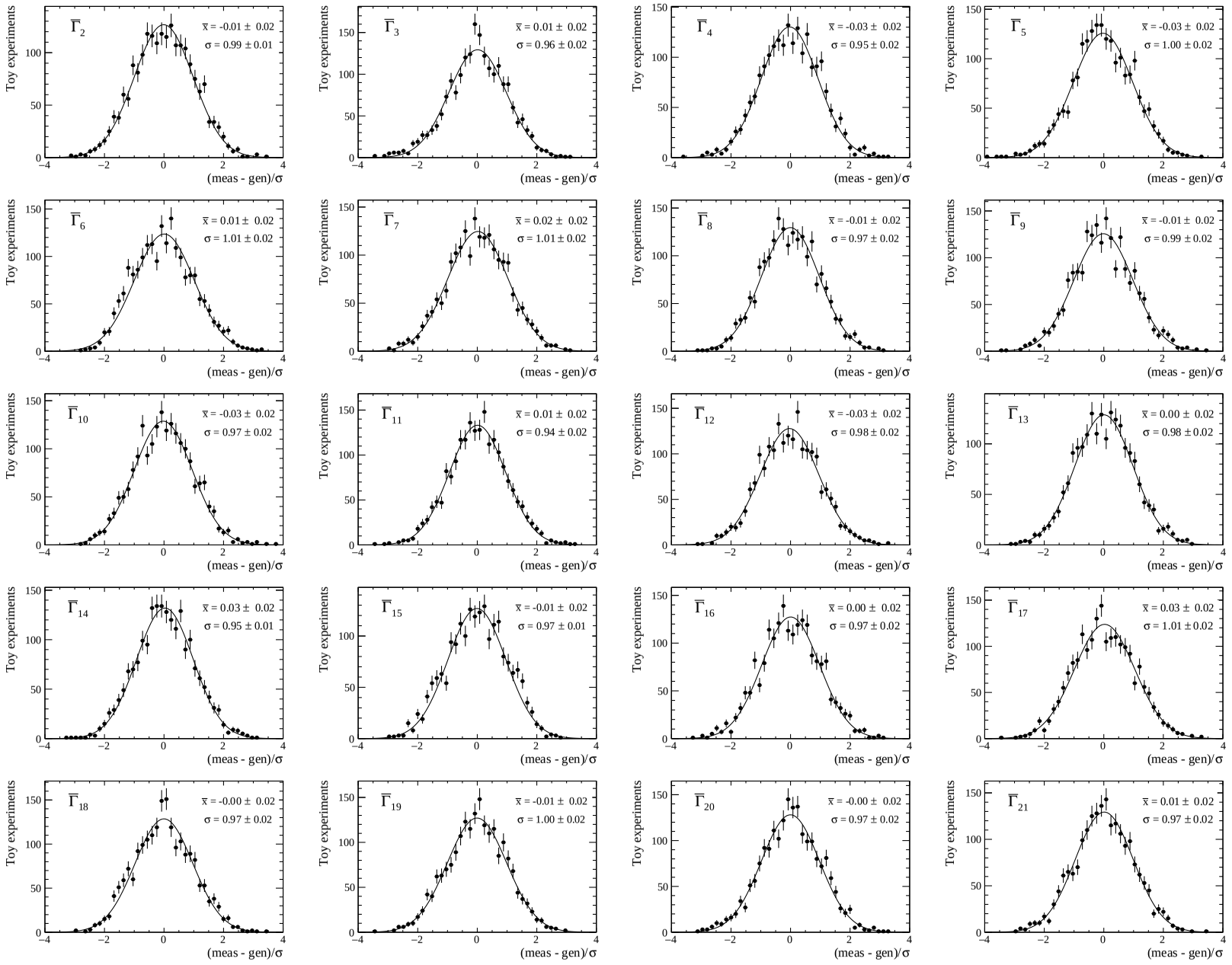


Figure 57: Pull distributions of the normalised moments, $\overline{\Gamma}_i$, for simulated datasets.

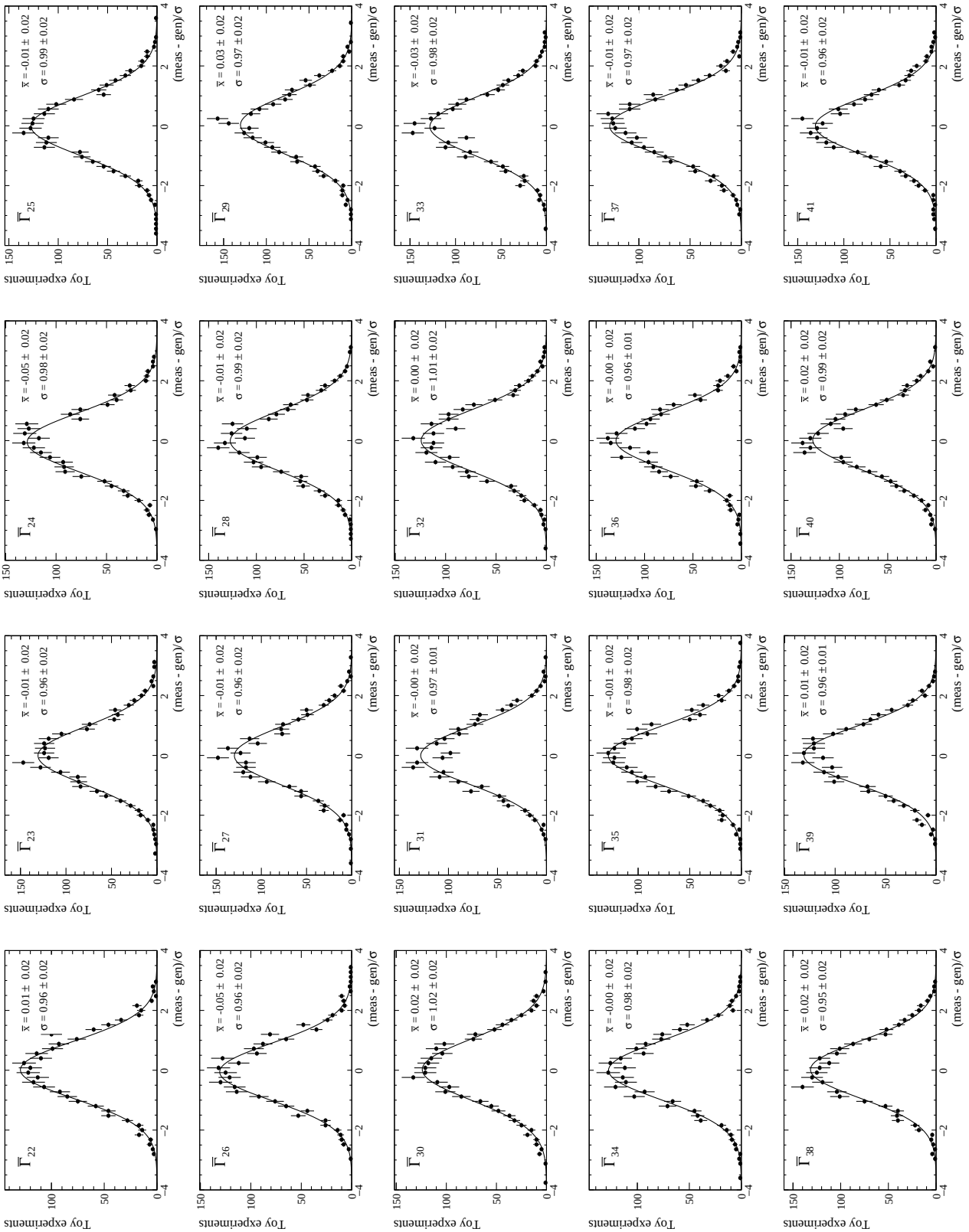


Figure 58: Pull distributions of the normalised moments, $\overline{\Gamma}_i$, for simulated datasets.

F Peaking background systematic

The effect of residual peaking background contributions is evaluated using pseudo-experiments where peaking background components are generated in addition to the signal and the combinatorial background. The peaking background candidates are first selected from data by isolating the decays using specific selections. In the following section, the method is shown explicity for the case of $B^0 \to J/\psi\, K^+\pi^-$ decays with a $\pi \leftrightarrow \mu$ swap.

As described in Sec. 7.3.5, candidates from $B^0 \to J/\psi\, K^+\pi^-$ decays can contribute a background if the π^- (K^+) is misidentified as a μ^- (μ^+) and the μ^- (μ^+) is misidentified as a π^- (K^+). For the case of $\mu^- \leftrightarrow \pi^-$ swaps, the invariant mass of the π^- and the μ^+, after assigning the π^- the μ mass hypothesis, should be consistent with the known J/ψ mass. This new mass hypothesis is denoted $m_{(\pi\to\mu)\mu}$.

Candidates are selected from data by requiring the $m_{(\pi\to\mu)\mu}$ invariant mass to be within $\pm 200\,\mathrm{MeV}/c^2$ of the known J/ψ mass. The $m_{(\pi\to\mu)\mu}$ distribution is fitted using a Gaussian to model the $B^0 \to J/\psi\, K^+\pi^-$ contribution and a second order Chebyshev polynomial to model the non-resonant contribution, as shown in Fig. 59. The *sPlot* technique [79] is used to isolate the $B^0 \to J/\psi\, K^+\pi^-$ contribution.

The distributions of the decay angles for $B^0 \to J/\psi\, K^+\pi^-$ decays with a $\pi \leftrightarrow \mu$ swap are shown in Fig. 60. A kernel estimator [91] is used to model the distributions, as shown by the blue histogram in Fig. 60. The resulting probability density function is used to generate the peaking background contributions for the pseudoexperiments.

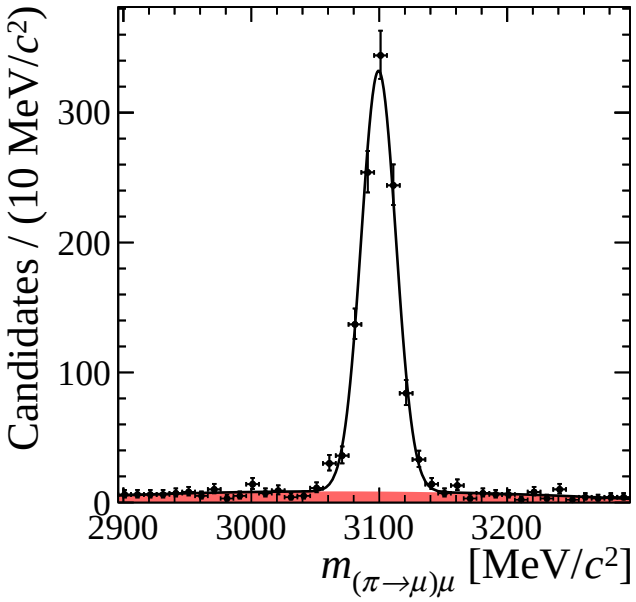


Figure 59: The $m_{(\pi\to\mu)\mu}$ distribution for $B^0 \to J/\psi\, K^+\pi^-$ candidates with a $\pi \leftrightarrow \mu$ swap.

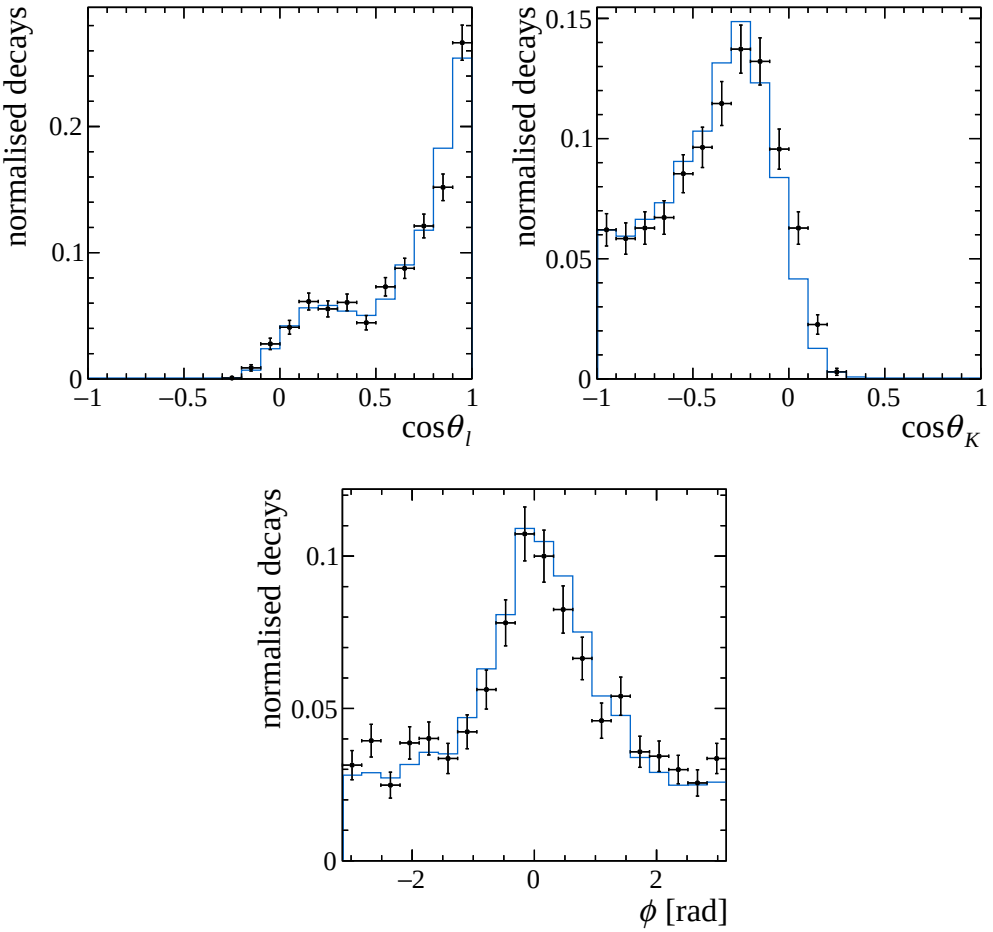


Figure 60: Distributions of each of the decay angles for $B^0 \to J/\psi K^+\pi^-$ decays with a $\pi \leftrightarrow \mu$ swap. The distributions derived from the kernel estimator are shown by the blue histograms.

Bibliography

[1] D. Griffiths, *Introduction to Elementary Particles*, John Wiley, 2008.

[2] F. Halzen and A. Martin, *Quarks and Leptons: an introductory course in modern particle physics*, Wiley, 1987.

[3] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics: a practical introduction*, CRC Press, 2013.

[4] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

[5] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[6] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[7] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[8] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285.

[9] E598, J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[10] SLAC-SP-017, J. E. Augustin *et al.*, *Discovery of a Narrow Resonance in e+ e- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406, [Adv. Exp. Phys.5,141(1976)].

[11] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[12] R. Gauld, F. Goertz, and U. Haisch, *An explicit Z'-boson explanation of the $B \to K^* \mu^+ \mu^-$ anomaly*, JHEP **01** (2014) 069, arXiv:1310.1082.

[13] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, Phys. Rev. **D89** (2014) 095033, arXiv:1403.1269.

[14] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Explaining $h \to \mu^\pm\tau^\mp$, $B \to K^*\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to Ke^+e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$*, Phys. Rev. Lett. **114** (2015) 151801, arXiv:1501.00993.

[15] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, arXiv:hep-ph/9806471.

[16] D. Becirevic and E. Schneider, *On transverse asymmetries in $B \to K^*\ell^+\ell^-$*, Nucl. Phys. **B854** (2012) 321, arXiv:1106.3283.

[17] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[18] LHCb Collaboration, A. A. Alves Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[19] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[20] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137, arXiv:hep-ph/0005110.

[21] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007. LHCb-TDR-001.

[22] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011. LHCb-TDR-005.

[23] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[24] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. LHCb-TDR-009.

[25] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029. LHCb-TDR-008.

[26] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024. LHCb-TDR-006.

[27] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.

[28] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.

[29] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[30] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036. LHCb-TDR-002.

[31] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010. LHCb-TDR-004.

[32] LHCb Collaboration, *LHCb muon system: addendum to the Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2003.

[33] LHCb Collaboration, *LHCb muon system: second addendum to the Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2005.

[34] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[35] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031. LHCb-TDR-010.

[36] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[37] LHCb Collaboration, V. V. Gligorov, *Performance and upgrade plans of the LHCb trigger system*, Nucl. Instrum. Meth. **A718** (2013) 26.

[38] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, Tech. Rep. LHCb-PUB-2011-016. CERN-LHCb-PUB-2011-016. LHCb-INT-2011-030, CERN, Geneva, Sep, 2011. LHCb-INT-2011-030.

[39] M. Williams *et al.*, *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.

[40] R. Aaij *et al.*, *Performance of the LHCb full real-time reconstruction and high-level trigger with 13 TeV data*, CERN-LHCb-DP-2016-002.

[41] R. Aaij *et al.*, *Tesla : an application for real-time data analysis in High Energy Physics*, arXiv:1604.05596.

[42] G. Dujany and B. Storaci, *Real-time alignment and calibration of the LHCb Detector in Run II*, LHCb-PROC-2015-011.

[43] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **10** (2015) 172, arXiv:1509.00771.

[44] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2016) 159, arXiv:1510.01707.

[45] LHCb collaboration, *Letter of Intent for the LHCb Upgrade*, CERN-LHCC-2011-001.

[46] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007. LHCb-TDR-012.

[47] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016. LHCb-TDR-016.

[48] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021. LHCb-TDR-013.

[49] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001. LHCb-TDR-015.

[50] E. Bowen, *Vertexing and Tracking Software at LHCb*, LHCb-PROC-2015-003.

[51] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, LHCb-PUB-2011-001.

[52] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, LHCb-2007-015.

[53] O. Callot and M. Schiller, *PatSeeding: A Standalone Track Reconstruction Algorithm*, LHCb-2008-042.

[54] M. Needham, *Performance of the Track Matching*, LHCb-2007-129.

[55] O. Callot, *Downstream Pattern Recognition*, LHCb-2007-026.

[56] S. Stahl, *Reconstruction of displaced tracks and measurement of K^0_S production rate in proton-proton collisions at $\sqrt{s}$ = 900 GeV at the LHCb experiment*, Masters thesis, Heidelberg University, Germany.

[57] O. Callot, M. Kucharczyk, and M. Witek, *VELO-TT track reconstruction*, LHCb-2007-010.

[58] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl.Instrum.Meth. A262 (1987) 444-450.

[59] J. van Tilburg, *Track simulation and reconstruction at LHCb*, PhD thesis, NIKHEF, Netherlands.

[60] E. Bowen and B. Storaci, *VeloUT tracking for the LHCb Upgrade*, LHCb-PUB-2013-023.

[61] E. Bowen, B. Storaci, and M. Tresch, *VeloTT tracking for LHCb Run II*, LHCb-PUB-2015-024.

[62] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region*, arXiv:1609.04736.

[63] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **108** (2012) 181806, arXiv:1112.3515.

[64] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$*, JHEP **08** (2013) 131, arXiv:1304.6325.

[65] LHCb collaboration, R. Aaij *et al.*, *Measurement of form-factor-independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013) 191801, arXiv:1308.1707.

[66] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442.

[67] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *On the impact of power corrections in the prediction of $B \to K^*\mu^+\mu^-$ observables*, JHEP **12** (2014) 125, arXiv:1407.8526.

[68] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **09** (2015) 179, arXiv:1506.08777.

[69] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular anaysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, JHEP **06** (2015) 115, arXiv:1503.07138.

[70] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, arXiv:1403.8044.

[71] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.

[72] W. Altmannshofer and D. M. Straub, *Implications of $b \to s$ measurements*, in *Proceedings, 50th Rencontres de Moriond Electroweak Interactions and Unified Theories: La Thuile, Italy, March 14-21, 2015*, pp. 333–338, 2015. arXiv:1503.06199.

[73] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *Global analysis of $b \to s\ell\ell$ anomalies*, JHEP **06** (2016) 092, arXiv:1510.04239.

[74] S. Sahoo and R. Mohanta, *Scalar leptoquarks and the rare B meson decays*, Phys. Rev. **D91** (2015), no. 9 094019, arXiv:1501.05193.

[75] S. Biswas, D. Chowdhury, S. Han, and S. J. Lee, *Explaining the lepton non-universality at the LHCb and CMS within a unified framework*, JHEP **02** (2015) 142, arXiv:1409.0882.

[76] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \to s\ell\ell$*, JHEP **02** (2015) 055, arXiv:1411.4773.

[77] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$?*, arXiv:1406.0566.

[78] C.-D. Lu and W. Wang, *Analysis of $B \to K_J^*(\to K\pi)\mu^+\mu^-$ in the higher kaon resonance region*, Phys. Rev. **D85** (2012) 034014, arXiv:1111.1513.

[79] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[80] B. Dey, *Angular analyses of exclusive $\overline{B} \to x\ell_1\ell_2$ with complex helicity amplitudes*, Phys. Rev. D **92** (2015) 033013.

[81] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $B \to K^*\mu^+\mu^-$ anomaly*, Phys. Rev. **D88** (2013) 074002, arXiv:1307.5683.

[82] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[83] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, Journal of Computer and System Sciences **55** (1997), no. 1 119 .

[84] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: Bounds for k-fold and progressive cross-validation*, in *Proceedings of the twelfth annual conference on Computational learning theory*, pp. 203–208, ACM, 1999.

[85] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays*, Phys. Rev. **D88** (2013) 052002, arXiv:1307.2782.

[86] Belle collaboration, K. Chilikin *et al.*, *Observation of a new charged charmoniumlike state in $\bar{B}^0 \to J/\psi K^-\pi^+$ decays*, Phys. Rev. **D90** (2014) 112009, arXiv:1408.6457.

[87] L. Demortier and L. Lyons, *Everything you always wanted to know about pulls*, CDF/ANAL/PUBLIC/5776.

[88] P. Ball and R. Zwicky, *$B_{d,s} \to \rho, \omega, K^*, \phi$ decay form factors from light-cone sum rules reexamined*, Phys. Rev. D **71** (2005) 014029.

[89] R.-H. Li, C.-D. Lü, W. Wang, and X.-X. Wang, *$b \to s$ transition form factors in the perturbative qcd approach*, Phys. Rev. D **79** (2009) 014013.

[90] W. Wang, *b to tensor meson form factors in the perturbative qcd approach*, Phys. Rev. D **83** (2011) 014008.

[91] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198, `arXiv:hep-ex/0011057`.